



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΦΥΣΙΚΗΣ
Τομέας Ε΄

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Παράλληλες Διεργασίες και Εφαρμογές σε
Επεξεργασία Εικόνων και Εικονορροών

Γεώργιος Δ. Γεωργής

ΑΘΗΝΑ
ΔΕΚΕΜΒΡΙΟΣ 2016



NATIONAL AND KAPODISTRIAN UNIVERSITY OF
ATHENS
SCHOOL OF SCIENCES
DEPARTMENT OF PHYSICS
Department E'

PhD THESIS

PARALLEL PROCESSES AND APPLICATIONS IN
IMAGE AND VIDEO PROCESSING

Georgios D. Georgis

ATHENS
DECEMBER 2016

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Παράλληλες Διεργασίες και Εφαρμογές σε Επεξεργασία Εικόνων και
Εικονορροών

Γεώργιος Δ. Γεωργής

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ

Διονύσιος Ρεΐσης, Αναπληρωτής Καθηγητής ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ ΠΑΡΑΚΟΛΟΥΘΗΣΗΣ:

Διονύσιος Ρεΐσης, Αναπληρωτής Καθηγητής ΕΚΠΑ

Άγγελος Κατσάγγελος, Καθηγητής Northwestern University

Εμμανουήλ Τσίλης, Επίκουρος Καθηγητής ΕΚΠΑ

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

(Υπογραφή)

(Υπογραφή)

Διονύσιος Ρεΐσης,
Αναπληρωτής Καθηγητής ΕΚΠΑ

Άγγελος Κατσάγγελος,
Καθηγητής Northwestern University

(Υπογραφή)

(Υπογραφή)

Εμμανουήλ Τσίλης,
Επίκουρος Καθηγητής ΕΚΠΑ

Δημήτριος Σούντρης,
Αναπληρωτής Καθηγητής ΕΜΠ

(Υπογραφή)

(Υπογραφή)

Ηλίας Μανωλάκος,
Καθηγητής ΕΚΠΑ

Ανδρέας Πολύδωρος,
Καθηγητής ΕΚΠΑ

(Υπογραφή)

Κωνσταντίνος Νικητόπουλος,
Επίκουρος Καθηγητής University of Surrey

21/12/2016 (Ημερομηνία εξέτασης)

PhD THESIS

PARALLEL PROCESSES AND APPLICATIONS IN IMAGE AND VIDEO
PROCESSING

Georgios D. Georgis

SUPERVISOR: Dionysios Reisis, Associate Professor NKUA

THREE-MEMBER ADVISORY COMMITTEE:

Dionysios Reisis, Associate Professor NKUA
Aggelos Katsaggelos, Professor Northwestern University
Emmanouil Tsilis, Assistant Professor NKUA

SEVEN-MEMBER EXAMINATION COMMITTEE

(Signature)

(Signature)

Dionysios Reisis,
Associate Professor NKUA

Aggelos Katsaggelos,
Professor Northwestern University

(Signature)

(Signature)

Emmanouil Tsilis,
Assistant Professor NKUA

Dimitrios Soudris,
Associate Professor NTUA

(Signature)

(Signature)

Elias Manolakos,
Professor NKUA

Andreas Polydoros,
Professor NKUA

(Signature)

Konstantinos Nikitopoulos,
Assistant Professor University of Surrey

21/12/2016 (Examination Date)

Περίληψη

Οι ταχέως εξελισσόμενες εφαρμογές επεξεργασίας εικονορροών εισαγάγουν απαιτήσεις υψηλής ποιότητας εικόνας, χαμηλού δυφιορρυθμού ή / και χαμηλού υπολογιστικού κόστους. Η τρέχουσα διατριβή μελετά το πεδίο της Υπερ-Ανάλυσης υπό το πρίσμα της μειωμένης πολυπλοκότητας, με σκοπό την παρουσίαση, την αποτίμηση επιδόσεων και την υλοποίηση παραλληλοποιήσιμων μεθόδων και τεχνικών. Επιπλέον, στοχεύει στην εφαρμογή αυτών των μεθόδων σε ένα γενικότερο πλαίσιο επεξεργασίας α) εικόνων και β) εικονορροών με συμπίεση αποδεκτικής-κβάντισης. Τέλος, μελετά τα απτά αποτελέσματα των τεχνικών αυτών σε πλατφόρμες επιτάχυνσης υλικού. Για το σκοπό αυτό, αρχικά μελετώνται τεχνικές υπερ-ανάλυσης χαμηλής πολυπλοκότητας προκειμένου να επιτευχθεί βελτίωση στη διαδικασία εκτίμησης κίνησης των κωδικοποιητών εικονορροών. Παρουσιάζονται τεχνικές παρεμβολής μειωμένου κόστους σε σύγκριση με το πρότυπο φίλτρο 6 δειγμάτων, με απώτερο στόχο την επιτάχυνση της παραγωγής των πλακιδίων κατά τη διαδικασία αναζήτησης. Η απόδοση των τεχνικών αυτών συγκρίνεται με αυτή των ευρέως διαδεδομένων τεχνικών ως προς την προκύπτουσα ποιότητα και το χρόνο επεξεργασίας. Η μελέτη βασίζεται στη χρήση ενός τυπικού αλγορίθμου κλασματικής εκτίμησης κίνησης για το πρότυπο H.264/AVC, και επομένως, ευνοεί το σχεδιασμό κωδικοποιητών βασισμένων στο πρότυπο αυτό. Κατόπιν, συνδυάζονται υπερσύγχρονες μέθοδοι υπερ-ανάλυσης και κωδικοποίησης με στόχο τη βελτίωση της αποτελεσματικότητας του κωδικοποιητή και την ελάττωση της πολυπλοκότητας. Η προτεινόμενη προσέγγιση βελτιώνει το γενικευμένο σχήμα συμπίεσης με αποδεκτική και κβάντιση, αφενός εισαγάγοντας τεχνικές υπερ-ανάλυσης χαμηλής πολυπλοκότητας για την υπερδειγματοληψία των δεδομένων στην πλευρά του αποκωδικοποιητή, αφετέρου εξερευνώντας και βελτιστοποιώντας τις διεργασίες αποδεκτικής / υπερδειγματοληψίας. Σε σχέση με το συμβατικό σχήμα κωδικοποίησης, το εμπλουτισμένο σχήμα επιτυγχάνει βελτίωση της ποιότητας και της πολυπλοκότητας, ενώ μπορεί εύκολα να παραμετροποιηθεί προκειμένου να υποστηρίξει οποιοδήποτε υφιστάμενο σχήμα κωδικοαποκωδίκευσης όπως το H.264 / AVC και το HEVC. Η προτεινόμενη προσέγγιση βασίζεται αρχικά στη μελέτη της παραμετροποίησης του γενικευμένου σχήματος με ευρέως διαδεδομένες τεχνικές μετασχηματισμού διαστάσεων, επιτυγχάνοντας βελτιωμένη τιμή κορυφής λόγου σήματος-θορύβου κατά 2 έως 4 dB σε σχέση με τα συμβατικά σχήματα. Εν συνεχεία εισάγεται ο αλγόριθμος *L-SEABI*, ένας νέος αλγόριθμος υπερ-ανάλυσης μέσω του οποίου αυξάνεται η τιμή του κρίσιμου δυφιορρυθμού στο επίπεδο των 10 Mb/s. Ο αλγόριθμος αποτιμάται και σε σύγκριση με άλλες λύσεις υπερδειγματοληψίας από τη βιβλιογραφία. Τα αποτελέσματα κα-

ταδεικνύουν βελτίωση της ποιότητας κατά 5 dB σε σχέση με τις απλές τεχνικές παρεμβολής και ελάττωση του χρόνου υπολογισμού κατά τρεις τάξεις μεγέθους σε σχέση με τις μεθόδους τεχνολογικής αιχμής. Τέλος, βασιζόμενοι στα χαρακτηριστικά του αλγορίθμου *L-SEABI* εισαγάγουμε τεχνικές παραλληλοποίησης για υλοποιήσεις σε μονάδες επεξεργασίας γραφικών και συστοιχίες επιτόπια προγραμματιζόμενων πυλών. Οι προτεινόμενες τεχνικές επιταχύνουν την ανακατασκευή περιεχομένου υπερ-υψηλής ευκρίνειας, επιτυγχάνοντας τριπλάσια απόδοση από τη συμβατική απαίτηση πραγματικού χρόνου σε μονάδες επεξεργασίας γραφικών μέσης απόδοσης και τουλάχιστον εννεαπλάσια απόδοση σε μονάδες επεξεργασίας γραφικών υψηλής απόδοσης. Αντίστοιχα αποτελέσματα της υλοποίησης σε FPGA καταδεικνύουν τετραπλασιασμό της απόδοσης των συμβατικών απαιτήσεων πραγματικού χρόνου σε μονάδες χαμηλής απόδοσης / κατανάλωσης και 69 φορές ταχύτερη απόδοση στη μονάδα Virtex 7 2000t.

ABSTRACT

Evolving video applications impose requirements for high image quality, low bitrate, and / or small computational cost. The current thesis involves the study of Super-Resolution techniques under the scope of reduced complexity aiming to present, evaluate and implement parallelizable methods and techniques. Moreover, it aims to apply these techniques to a more generic framework of processing a) images and b) video sequences with decimation and quantization. Finally, it studies these methods' implementation feasibility on hardware acceleration platforms. To that end, we initially study low-complexity image super-resolution techniques for improving the motion estimation process of video encoding. Aiming at speeding up the generation of candidate blocks during the computationally intensive search algorithm, we present interpolation techniques with reduced cost compared to standard 6-tap filtering procedures. Furthermore, we compare their performance to that of commonly used half-pixel interpolation techniques with respect to the resulting image quality and the processing time. The research has been based on using a typical fractional motion estimation algorithm preceding the processing of the H.264 / AVC standard motion compensation, and thus, the research benefits the design of H.264 / AVC encoders. Next, we combine state-of-the-art coding and super-resolution (SR) techniques to improve video compression both in terms of coding efficiency and complexity. The proposed approach improves a generic decimation-quantization compression scheme by introducing low complexity single-image SR techniques for rescaling the data at the decoder side and by jointly exploring / optimizing the downsampling / upsampling processes. The enhanced scheme achieves improvement of the quality and system's complexity compared with conventional codecs and can be easily modified to meet various diverse requirements, such as effectively supporting any off-the-shelf video codec, for instance H.264 / Advanced Video Coding or High Efficiency Video Coding. Our approach builds on studying the generic scheme's parameterization with common rescaling techniques to achieve 2.4-dB peak signal-to-noise ratio (PSNR) quality improvement at low-bitrates compared with the conventional codecs and proposes *L-SEABI*, a novel SR algorithm to advance the critical bitrate at the level of 10 Mb/s. The evaluation of the SR algorithm includes the comparison of its performance to other image rescaling solutions of the literature. The results show quality improvement by 5 dB PSNR over straight-forward interpolation techniques and computational time reduction by three orders of magnitude when compared with the highly involved methods of the

field. Finally, the study builds up on the characteristics of the *L-SEABI* SR method to introduce parallelization techniques for GPUs and FPGAs. The proposed techniques accelerate GPU reconstruction of ultra-high definition content, by achieving three ($3\times$) times faster than the real-time performance on mid-range and previous generation devices and at least nine times ($9\times$) faster than the real-time performance on high-end GPUs. The FPGA design leads to a scalable architecture performing four ($4\times$) times faster than the real-time on low-end Xilinx Virtex 5 devices and 69 times ($69\times$) faster than the real-time on the Virtex 2000t.

keywords: super-resolution; image / video processing; video compression; graphics processing units; field-programmable gate arrays

λέξεις-κλειδιά: υπερ-ανάλυση; επεξεργασία εικόνων/εικονορροών; συμπίεση εικονορροών; μονάδες επεξεργασίας γραφικών; συστοιχίες επιτόπια προγραμματιζόμενων πυλών

Αριθμός Μητρώου: 2009505

Δ.Δ.Φ.Ε. Ηλεκτρονικής, Υπολογιστών, Τηλεπικοινωνιών και Αυτοματισμού

Στην οικογένειά μου

Ευχαριστίες

Από όλη αυτή την πορεία θα ήθελα πρωταρχικά να ευχαριστήσω τον "καπετάνιο" και επιβλέποντα της διατριβής αναπληρωτή καθηγητή κ. Δ. Ρεΐση, για την πολυεπίπεδη βοήθειά του, για τον πατρικό του χαρακτήρα και γιατί ήταν πάντα παρών προκειμένου να διατηρηθεί η πολύ μικρή ομάδα του εργαστηρίου αργαγής. Ιδιαίτερες ευχαριστίες οφείλω στο συνάδελφο Δρ. Γεώργιο Λεντάρη, για τη βοήθεια, την καθοδήγηση και τη συνεργασία του σε όλη τη διάρκεια αλλά και μετά την περάτωση αυτού του εγχειρήματος. Δε θα ήταν υπερβολή να γράψω ότι η στάση του απέναντι στην επιστήμη αποτέλεσε πρότυπο και διαμόρφωσε καθοριστικό ρόλο ώστε να αποφασίσω να ξεκινήσω και να ολοκληρώσω τις διδακτορικές μου σπουδές. Ευχαριστώ όλους τους συνεργάτες και φίλους σε όλα αυτά τα χρόνια και ειδικότερα: τον Νίκο, τον Γιάννη, τον Παναγιώτη, τον Κώστα, τον Ορέστη, τον Χάρη και τον Γιώργο. Είμαι ευγνώμων για την "παρέα του εργαστηρίου" ήτοι τον Νίκο, τον Κώστα, τον Γιώργο, τον Χάρη, τον Παναγιώτη, την Αφροδίτη, τη Φωτεινή, και φυσικά τους καλούς μου φίλους, τον Βαγγέλη, τον Αποστόλη, τον Φάνη και τον Σταύρο, αφού χάρις στην παρουσία τους απέκτησα το προνόμιο να μιλώ για ανέμελες στιγμές στα πολύ δύσκολα αυτά χρόνια. Οφείλω να ευχαριστήσω την επιστημονική κοινότητα, για την καίρια καθοδήγηση και τον υποδειγματικό της χαρακτήρα, από τον καθηγητή κ. Φραντζεσκάκη, την επιτροπή εξέτασης του διδακτορικού, το συνεργάτη στο νέο μου περιβάλλον επίκουρο καθηγητή κ. Νικητόπουλο, έως και τον ανώνυμο συνάδελφο που εργάζεται για να παράξει γνώση μέσα σε όλες αυτές τις αντιξοότητες της ελληνικής πραγματικότητας.

Ξεχωριστές ευχαριστίες ανήκουν στην Οικογένειά μου, τους Γονείς μου, τον Αδερφό μου και την αγαπημένη μου Αθανασία. Χωρίς την αμέριστη συμπαράσταση και τις θυσίες τους, δε θα μπορούσα ποτέ να επιτύχω αυτό το στόχο. Κλείνοντας, θα ήθελα να ευχαριστήσω και τη φίλη μου τη Ρούλα, για την υπομονή, τη στήριξη και τη βοήθειά της στο τελευταίο μέρος του ταξιδιού. Ένα συναρπαστικό και τρικυμιώδες ταξίδι, το οποίο έφθασε όχι στο τέλος του, αλλά στον πρώτο του προορισμό.

Λίστα Δημοσιεύσεων

- [1] C. Husmann, G. Georgis, K. Nikitopoulos, and K. Jamieson. “*FlexCore: Massively Parallel and Flexible Processing for Large MIMO Access Points*”. In Networked Systems Design and Implementation (NSDI), 2017 14th USENIX symposium on, Mar 2017 (Accepted).
- [2] Georgios Georgis, George Lentaris, and Dionysios Reisis. “*Acceleration techniques and evaluation on multi-core cpu, gpu and fpga for image processing and super-resolution*”. Journal of Real-Time Image Processing, pages 1–28, July 2016.
- [3] G. Georgis, G. Lentaris, and D. Reisis. “*Reduced complexity superresolution for low-bitrate video compression*”. IEEE Transactions on Circuits and Systems for Video Technology, 26(2):332–345, Feb 2016.
- [4] G. Georgis, G. Menoutis, D. Reisis, K. Tsakalis, and A. B. Shafique. “*Towards real-time neuronal connectivity assessment: A scalable pipelined parallel generalized partial directed coherence engine*”. In 2015 IEEE International Conference on Electronics, Circuits, and Systems (ICECS), pages 13–16, Dec 2015.
- [5] G. Menoutis, A. Foteas, N. Liakopoulos, G. Georgis, D. Reisis, and G. Synnefakis. “*A configurable transmitter architecture organization for xg-pon olt/onu/ont network elements*”. In 2015 IEEE International Conference on Electronics, Circuits, and Systems (ICECS), pages 673–676, Dec 2015.
- [6] G. Georgis, C. Tzeranis, D. Reisis, and G. Synnefakis. “*XG-PON optical network unit downstream fec design based on truncated reed-solomon code*”. In Electronics, Circuits and Systems (ICECS), 2014 21st IEEE International Conference on, pages 782–785, Dec 2014.
- [7] G. Georgis, D. Reisis, P. Skordilakis, K. Tsakalis, A. B. Shafique, G. Chatzikonstantis, and G. Lentaris. “*Neuronal connectivity assessment for epileptic seizure prevention: Parallelizing the generalized partial directed coherence on many-core platforms*”. In Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS XIV), 2014 International Conference on, pages 359–366, July 2014.
- [8] G. Georgis, C. Tzeranis, D. Reisis, and G. Synnefakis. “*Fpga design of the decoding functions in the physical layer adaptation subsystem of the xg-pon optical network unit/terminal*”. In Ph.D. Research in Microelectronics and Electronics (PRIME), 2014 10th Conference on, pages 1–4, June 2014.
- [9] G. Georgis, G. Lentaris, and D. Reisis. “*Single-image super-resolution using low-complexity adaptive iterative back-projection*”. In Digital Signal Processing. 18th IEEE International Conference on, pages 152–157, 2013.

- [10] G. Georgis, G. Lentaris, and D. Reisis. Design and Architectures for Digital Signal Processing, chapter “*Low Complexity Interpolation Filters for Motion Estimation and Application to the H.264 Encoders*”. Intech Open Access, 2013.
- [11] K. Manolopoulos, A. Belias, G. Georgis, D. Reisis, and E. G. Anasontzis. “*Signal processing for deep-sea observatories with reconfigurable hardware*”. In Electronics, Circuits and Systems (ICECS), 2012 19th IEEE International Conference on, pages 81–84, Dec 2012.
- [12] G. Georgis, G. Lentaris, and D. Reisis. “*Study of interpolation filters for motion estimation with application in h.264/avc encoders*”. In Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on, pages 9–12, 2011.

Περιεχόμενα

Περίληψη	i
Abstract	ii
Κατάλογος σχημάτων	5
Κατάλογος πινάκων	9
1 Εισαγωγή	13
1.1 Συμπίεση εικονορροών: εκτίμηση κίνησης	13
1.2 Συμπίεση εικονορροών με αποδεκίαση	14
1.3 Υπερ-Ανάλυση	15
1.4 Στόχοι και διάρθρωση της διατριβής	16
2 Πρόβλημα και Σχετική Βιβλιογραφία	19
2.1 Συμπίεση με αποδεκίαση και κβάντιση	19
2.2 Αλγόριθμοι Υπερ-Ανάλυσης	21
2.3 Επιτάχυνση αλγορίθμων σε υλικό	23
2.3.1 Επιτάχυνση σε Μονάδες Επεξεργασίας Γραφικών	23
2.3.2 Επιτάχυνση με Συστοιχίες Επιτόπια Προγραμματιζόμενων Πυλών (ΣΕΠΠ, Field-Programmable Gate Arrays - FPGAs)	23
2.3.3 Συγκριτικές μελέτες επιτάχυνσης σε Μονάδες Επεξεργα- σίας Γραφικών και Συστοιχίες Επιτόπια Προγραμματιζό- μενων Πυλών	24
3 Τεχνικές Φίλτρων Παρεμβολής για Εκτίμηση Κίνησης με Εφαρμογή σε Κωδικοποιητές H.264/AVC	27
3.1 Τεχνικές Παρεμβολής	28
3.1.1 Διγραμμική Παρεμβολή	28

3.1.2	Δικυβική Παρεμβολή	28
3.1.3	Lanczos Παρεμβολή	30
3.1.4	Τριγωνισμός Εξαρτώμενος από τα Δεδομένα εισόδου (ΤΕΔ, Data-Dependent Triangulation - DDT)	30
3.1.5	Αχμοστρεφής τεχνική CrossHD	31
3.1.6	Αχμοστρεφής τεχνική CxScale	32
3.2	Αξιολόγηση τεχνικών	33
4	Υπερ-ανάλυση από Μία Εικόνα με τη χρήση Χαμηλής Πολυπλοκότητας Προσαρμοστικής Επαναλαμβανόμενης Οπισθοπροβολής	39
4.1	Μοντέλο παρατήρησης εικόνων και ΕΟΠ από μία εικόνα	39
4.2	Προτεινόμενος αλγόριθμος ΥΑ	40
4.2.1	Εξερεύνηση τεχνικών παρεμβολής	41
4.2.2	Φίλτρο αποδεκάτισης και πυρήνας οπισθοπροβολής	45
4.3	Αξιολόγηση ποιότητας και πολυπλοκότητας	46
5	Μειωμένης Πολυπλοκότητας ΥΑ για Συμπίεση Εικονορροών Χαμηλού Δυφιορρυθμού	55
5.1	Προτεινόμενο σχήμα συμπίεσης	56
5.2	Μελέτη και ανάπτυξη αλγορίθμων ΥΑ	60
5.2.1	Προτεινόμενος αλγόριθμος ΥΑ: L-SEABI	62
5.2.2	Αξιολόγηση αλγορίθμων ΥΑ και του L-SEABI	68
5.3	ΥΑ μέσω <i>L-SEABI</i> στο προτεινόμενο σχήμα συμπίεσης	76
5.3.1	Αξιολόγηση δυφιορρυθμού-παραμόρφωσης	76
5.3.2	Αξιολόγηση πολυπλοκότητας στο σχήμα συμπίεσης	79
6	Τεχνικές Παραλληλοποίησης και Υλοποίησης σε Υλικό	83
6.1	Αλγόριθμος SIL-SEABI: επιτάχυνση σε ΜΕΓ	84
6.1.1	Στρατηγική υλοποίησης: από τη CPU στη ΜΕΓ	84
6.1.2	Λεπτομέρειες υλοποίησης σε ΜΕΓ και βελτιστοποιήσεις	85
6.2	Επιτάχυνση του L-SEABI σε FPGA	99
6.2.1	Σχεδίαση παράλληλης αρχιτεκτονικής για τον αλγόριθμο SIL-SEABI	99

6.2.2	Εξερεύνηση χώρου σχεδίασης και αποτελέσματα υλοποίησης	103
6.3	Σύγκριση υλοποιήσεων σε CPU, GPU και FPGA	106
6.3.1	L-SEABI: αποτελέσματα ποιότητας	106
6.3.2	Αποτελέσματα επιτάχυνσης L-SEABI: ανάλυση υλοποίησης σε CPU	106
6.3.3	Αποτελέσματα επιτάχυνσης L-SEABI: ανάλυση υλοποίησης σε MEG	110
6.3.4	Αποτελέσματα επιτάχυνσης L-SEABI: ανάλυση υλοποίησης σε FPGA	110
6.3.5	Συγκρίσεις πλατφόρμας υλοποίησης: ισχύς και απόδοση	111
6.3.6	Συγκρίσεις πλατφόρμας υλοποίησης: κόστος ανάπτυξης	114
7	Πλεονεκτήματα και εφαρμογές των προτεινόμενων τεχνικών	117
7.1	Ο αλγόριθμος <i>L-SEABI</i> ως μηχανισμός βελτίωσης των υπερσύγχρονων αλγορίθμων ΥΑ	117
7.2	Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγορίθμους στερεοσκοπικής αντιστοίχισης	125
7.2.1	Υλοποίηση ανομοιότητας σε MEG	126
7.2.2	Υλοποίηση ανομοιότητας σε FPGA	129
8	Συμπεράσματα	133
	Βιβλιογραφία	137

Κατάλογος σχημάτων

3.1	Πλέγμα εικονοστοιχείων εικόνας και απεικόνιση θέσεων υποεικονοστοιχείων με μεγέθυνση περιοχής διαστάσεων 1×1 (δεξιά). Ο συμβολισμός επελέγη ώστε να διευκολύνει την περιγραφή των φίλτρων.	29
4.1	Ο προτεινόμενος αλγόριθμος ΥΑ.	42
4.2	Υποκειμενική σύγκριση αλγορίθμων ΥΑ: 4.2α') εικόνα αναφοράς, 4.2β') δικυβική, 4.2γ') Matlab IBP Lanczos3, 4.2δ') προτεινόμενη.	50
4.3	Απόδοση ΥΑ (Y-PSNR) συναρτήσει του παράγοντα υπερ-δειγματοληψίας: σύγκριση 3 αλγορίθμων με εικόνα αναφοράς QCIF.	51
4.4	Απόδοση ΥΑ (Y-PSNR) συναρτήσει του παράγοντα υπερ-δειγματοληψίας: σύγκριση 3 αλγορίθμων με εικόνα αναφοράς 1080p.	52
5.1	Συμβατικό και προτεινόμενο σχήμα κωδικοποίησης.	57
5.2	Επιδόσεις του προτεινόμενου σχήματος συμπίεσης με τη χρήση διακριτών προτύπων κωδικοποίησης και φίλτρων παρεμβολής (25fps).	59
5.3	Τα βήματα 1-3 του αλγορίθμου Υπερ-Ανάλυσης <i>L-SEABI</i>	66
5.4	Το 4ο βήμα του αλγορίθμου <i>L-SEABI</i>	67
5.5	Αντιπαράθεση των πιο διακεκριμένων ποιοτικά αλγορίθμων κάθε κατηγορίας για υποκειμενική σύγκριση, $f = 2$	73
5.6	Από κοινού αποτίμηση ποιότητας/πολυπλοκότητας αλγορίθμων ΥΑ.	75
5.7	Βελτίωση Δυφιορρυθμού-Παραμόρφωσης στο AVC μέσω του <i>L-SEABI</i>	77
5.8	Βελτίωση Δυφιορρυθμού-Παραμόρφωσης στο HEVC μέσω του <i>L-SEABI</i>	78
5.9	Πολυπλοκότητα συμβατικής και προτεινόμενης συμπίεσης.	81
6.1	Οι διακριτές διεργασίες του αλγορίθμου <i>L-SEABI</i>	84
6.2	Κατατομή χρόνου διεργασιών του αλγορίθμου <i>SIL-SEABI</i> για μονονηματική εκτέλεση σε επεξεργαστή γενικού σκοπού.	85

6.3	Ψευδοκώδικας υλοποίησης <i>L-SEABI</i> στη GPU: ιεραρχία πυρήνων.	87
6.4	<i>L-SEABI</i> : Κατατομή χρόνου εκτέλεσης ΜΕΓ για μία επανάληψη.	89
6.5	Κώδικας υπολογισμού εικονοστοιχείων σε HO θέσεις εντός του πυρήνα <i>mcxHDc</i> : υπολογισμός χωρίς νηματικές αποκλίσεις. . . .	94
6.6	Ανάλυση βελτιστοποιήσεων σε ΜΕΓ για τον αλγόριθμο <i>SIL-SEABI</i> : μέση τιμή επιτάχυνσης ανά επεξεργαστικό πυρήνα για όλες τις αναλύσεις/περιεχόμενο εικόνων εν συγκρίσει με τη βασική υλοποίηση.	95
6.7	Χρόνος εκτέλεσης <i>SIL-SEABI</i> σε ΜΕΓ.	98
6.8	Προτεινόμενη αρχιτεκτονική διάταξη ανάγνωσης/επεξεργασίας σειριακού σε παράλληλο για τη δισδιάστατη σάρωση πλεγματοειδούς εικόνας με ολισθαίνον παράθυρο $N \times N$ εικονοστοιχείων.	101
6.9	Η αρχιτεκτονική μίας μηχανής ΥΑ του αλγορίθμου <i>L-SEABI</i> . Διακρίνονται οι σωληνώσεις οι οποίες περιλαμβάνουν μνήμες μετατροπής σειριακού σε παράλληλο (B) και το επεξεργαστικό σκέλος των παράλληλων μονάδων (PUs). Ο ρυθμός επεξεργασίας ποικίλλει από ένα έως τέσσερα δείγματα ανά κύκλο αναλόγως της χωρικής ανάλυσης της εικόνας.	102
6.10	Συνοπτική εικόνα της απόδοσης του αλγορίθμου <i>SIL-SEABI</i> ανά πλατφόρμα επιτάχυνσης και ανάλυση αποτελέσματος: Απαιτούμενος χρόνος υπερ-δειγματοληψίας εικόνας ΧΕ κατά παράγοντα $f = 2$.	109
6.11	Χρόνος και Ισχύς κατά τη διεργασία ΥΑ μέσω του αλγορίθμου <i>SIL-SEABI</i> . Στο γράφημα περιλαμβάνονται διάφορα ολοκληρωμένα κυκλώματα μαζί με το αντίστοιχο κόστος κτήσης τους (σε USD) (ο συμβολισμός εξηγείται στο κυρίως κείμενο).	112
6.12	Συνδυαστικός δείκτης απόδοσης-κόστους του αλγορίθμου <i>SIL-SEABI</i> σε διάφορες πλατφόρμες υλοποίησης.	115
7.1	Υποκειμενική σύγκριση ποιότητας των μεθόδων [24], [98]: τυπική επεξεργασία έναντι εμπλουτισμένης με τον αλγόριθμο <i>SIL-SEABI</i> ($f = 2$).	124
7.2	Αλγόριθμος <i>Ανομοιότητας</i> στην πλατφόρμα CUDA: ροή επεξεργασίας στη ΜΕΓ και πυρήνες επεξεργασίας.	126
7.3	Χρόνος εκτέλεσης υλοποιήσεων αλγορίθμου ανομοιότητας σε ΜΕΓ.	128
7.4	Χρόνος και Ισχύς κατά τη διεργασία ΥΑ μέσω του αλγορίθμου <i>SIL-SEABI</i> . Στο γράφημα περιλαμβάνονται διάφορα ολοκληρωμένα κυκλώματα μαζί με το αντίστοιχο κόστος κτήσης τους (σε USD) (ο συμβολισμός εξηγείται στο κυρίως κείμενο).	132

Κατάλογος πινάκων

3.1	Μετρική PSNR για το πρότυπο φίλτρο του H.264/AVC και ΔPSNR για τις υπόλοιπες τεχνικές όταν η ΕΚ αφορά αποκλειστικά HO+HK θέσεις (με αντιστάθμιση προτύπου H.264).	34
3.2	Μετρική PSNR για το πρότυπο φίλτρο του H.264/AVC και ΔPSNR για τις υπόλοιπες τεχνικές όταν η ΕΚ αφορά αποκλειστικά ΗΔ θέσεις (με αντιστάθμιση προτύπου H.264).	36
3.3	Μετρήσεις Ποιότητας και Χρόνου υπολογισμού κατά την Εκτίμηση Κίνησης σε HO+HK+ΗΔ θέσεις.	37
4.1	Μετρική PSNR (dB) για τη Δικυβική παρεμβολή και ΔPSNR για τις υπόλοιπες μεθόδους όταν πραγματοποιείται υπερ-δειγματοληψία κατά $f = 4$	48
4.2	Μετρική MSSIM για τη Δικυβική παρεμβολή και ΔMSSIM για τις υπόλοιπες μεθόδους όταν πραγματοποιείται υπερ-δειγματοληψία κατά $f = 4$	49
4.3	Απαιτούμενες επαναλήψεις για τη σύγκλιση σχημάτων ΕΟΠ.	52
5.1	Αντικειμενική σύγκριση ποιότητας του προτεινόμενου αλγορίθμου <i>L-SEABI</i> έναντι της Δικυβικής παρεμβολής και υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f = 2$). Όλοι οι αλγόριθμοι έχουν εκτελεστεί με τις προτεινόμενες από τους συγγραφείς τους παραμέτρους. Η χαμηλότερη τιμή της μετρικής BRISQUE αντιστοιχεί σε υψηλότερη ποιότητα εικόνας.	71
5.2	Σύγκριση πολυπλοκότητας: Μέσος χρόνος υπερ-δειγματοληψίας (sec) μεθόδων ΥΑ ($f = 2$) και μέση τιμή ανά εικονοστοιχείο (μ.τ./ε.σ.).	74
5.3	Επιδόσεις Σχήματος Συμπίεσης με ΥΑ.	79
6.1	Τύπος εντολής PTX για τη διεργασία της διαχωρίσιμης συνέλιξης, αριθμός εκτελέσεων και ρυθμοαπόδοση ανά αριθμητική/επεξεργαστή γραφικών.	92

6.2	Κόστος πόρων διάταξης FPGA έναντι χρόνου εκτέλεσης (στις διατάξεις εισαγάγεται μία εικόνα διαστάσεων 1920×1080 και επιλέγονται διάφοροι βαθμοί παραλληλίας για παράγοντα υπερδευγματοληψίας $f = 2$)	104
6.3	Επιδόσεις Ποιότητας των υλοποιήσεων του αλγορίθμου <i>SIL-SEABI</i> σε CPU, GPU και FPGA.	107
6.4	Επιδόσεις Επιτάχυνσης του αλγορίθμου <i>SIL-SEABI</i> σε CPU, GPU και FPGA.	108
6.5	Ενεργειακή αποτίμηση των υλοποιήσεων του αλγορίθμου <i>SIL-SEABI</i> σε CPU, GPU και FPGA.	114
7.1	Αντικειμενική σύγκριση ποιότητας μεταξύ του αλγορίθμου <i>L-SEABI</i> της εκτέλεσης μίας επανάληψης αυτού, της φάσης δόμησης αυτού και των μεθόδων ΥΑ <i>NARM</i> και της ΥΑ μέσω παραμορφώσιμων πλακιδίων (παράγοντας υπερ-δευγματοληψίας $f = 2$). Όλοι οι αλγόριθμοι εκτελούνται με τις προτεινόμενες παραμέτρους από τους αντίστοιχους συγγραφείς. Χαμηλότερες τιμές της μετρικής BRISQUE υποδηλώνουν υψηλότερη αντικειμενική ποιότητα εικόνων.	118
7.2	Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δευγματοληψίας $f = 2$) όταν χρησιμοποιούνται οι αλγόριθμοι <i>L-SEABI</i> (a), <i>SIL-SEABI</i> (b) και <i>L-SEAI</i> (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική MSSIM).	119
7.3	Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δευγματοληψίας $f = 2$) όταν χρησιμοποιούνται οι αλγόριθμοι <i>L-SEABI</i> (a), <i>SIL-SEABI</i> (b) και <i>L-SEAI</i> (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική MSSIM).	120
7.4	Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δευγματοληψίας $f = 2$) όταν χρησιμοποιούνται οι αλγόριθμοι <i>L-SEABI</i> (a), <i>SIL-SEABI</i> (b) και <i>L-SEAI</i> (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική BRISQUE). Χαμηλότερες τιμές διαφοράς ΔBRISQUE υποδηλώνουν υψηλότερη αντικειμενική ποιότητα εικόνων.	121

7.5	Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f=2$) όταν χρησιμοποιούνται οι αλγόριθμοι <i>L-SEABI</i> (a), <i>SIL-SEABI</i> (b) και <i>L-SEAI</i> (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική BRISQUE). Χαμηλότερες τιμές διαφοράς ΔBRISQUE υποδηλώνουν υψηλότερη αντικειμενική ποιότητα εικόνων.	122
7.6	Ρυθμός πλαισίων ανομοιότητας σε ΜΕΓ.	129

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Η τελευταία δεκαετία καθορίστηκε από μία διαρκώς αυξημένη ζήτηση σε φορητές υπολογιστικές συσκευές, όπως φορητούς προσωπικούς υπολογιστές, έξυπνα τηλέφωνα και υπολογιστές-πινακίδια (tablets). Με την παγίωση αυτής της φορητότητας ενισχύθηκαν και οι ερευνητικές προσπάθειες σε πεδία όπως η κινητή υπολογιστική (mobile computing) και οι ασύρματες τηλεπικοινωνίες, ειδικά στον τομέα της μετάδοσης και επεξεργασίας οπτικοποιημένων πληροφοριών. Μεγάλο εύρος οπτικοακουστικών εφαρμογών χρειάζεται να συνδυάζει με αποδοτικό τρόπο χαρακτηριστικά κατά κανόνα αντικρουόμενα, όπως χαμηλό δυφιόρρυθμό (bitrate), μειωμένη πολυπλοκότητα και ταυτόχρονα υψηλό επίπεδο ποιότητας εικόνας και παρεχόμενων υπηρεσιών.

Όλα τα βήματα εξέλιξης αναφορικά με την ποιότητα των οπτικών εφαρμογών βασίζονται στη βελτίωση της αποδοτικότητας των προτύπων συμπίεσης ως προς το δίπολο δυφιόρρυθμού-παραμόρφωσης (rate-distortion). Προς αυτή την κατεύθυνση κινήθηκαν διαδοχικές γενιές προτύπων κωδικοποίησης με αξιολογικά παραδείγματα αυτά των προτύπων MPEG-2 και H.263 στη δεκαετία του 1990, τα MPEG-4 και H.264/AVC στη δεκαετία του 2000 [64] και το πρόσφατο HEVC (H.265) [83]. Παρά το γεγονός ότι κάθε νέα γενιά κωδικοαποκωδικοποιητών (codecs) διατηρεί την ποιότητα της προηγούμενης γενιάς απαιτώντας μόλις το ήμισυ του δυφιόρρυθμού, για να το επιτύχει αυξάνει σημαντικά το υπολογιστικό κόστος της συμπίεσης. Σε κάθε πρότυπο κωδικοαποκωδικοποίησης, ο κωδικοποιητής περιλαμβάνει το μεγαλύτερο μέρος του υπολογιστικού φόρτου, τη στιγμή που ο αποκωδικοποιητής χαρακτηρίζεται από 2 έως 4 τάξεις μεγέθους χαμηλότερη πολυπλοκότητα.

1.1 Συμπίεση εικονορροών: εκτίμηση κίνησης

Η διεργασία εκτίμησης κίνησης (EK, Motion Estimation - ME), αποτελεί γενικά το σημαντικότερο υπολογιστικά σκέλος ενός κωδικοποιητή. Περιλαμβάνει την ανάλυση κίνησης της εικονορροής σε επίπεδο πλακιδίων (δισδιάστατων τμημάτων) μεταξύ διαδοχικών πλαισίων (frames), με σκοπό την εύρεση του βέλτιστου ταιριάσματος κάθε πλακιδίου στο χώρο αναφοράς των πλαισίων της εικονορροής. Το ταιρίασμα επιτυγχάνεται με μεθόδους μέτρησης της απομένουσας

ενέργειας μέσω του υπολογισμού των διαφορών στις τιμές φωτεινότητας/χρώματος των πλακιδίων. Την ΕΚ ακολουθεί η αντιστάθμιση κίνησης, η οποία αξιοποιεί τα υπολογισμένα ανύσματα κίνησης και τη διαφορά φωτεινότητας/χρώματος των πλακιδίων στο σημείο ταιριάσματος, προκειμένου να επανασυνθέσει το κάθε πλακίδιο - και κατά συνέπεια το πλήρες πλαίσιο. Συχνά, η μετακίνηση των πλακιδίων γίνεται κατά μη ακέραιο αριθμό θέσεων. Σε αυτή την περίπτωση, διενεργείται κλασματική εκτίμηση κίνησης (Fractional Motion Estimation - FME) η οποία διευκολύνει την αφαίρεση πλεονασμών στο πεδίο του χρόνου και βελτιώνει την ποιότητα των εικονορροών κατά 2-4 dB με βάση τη μετρική της κορυφής λόγου σήματος-θορύβου (Peak Signal-to-Noise-Ratio, PSNR) [37] [16]. Ωστόσο, η αύξηση της πολυπλοκότητας την οποία συνεπάγεται η κλασματική εκτίμηση κίνησης, αυξάνει ακόμη περισσότερο τον υπολογιστικό φόρτο του κωδικοποιητή με αποτέλεσμα τον περιορισμό ή την αδυναμία εφαρμογής σε συνθήκες επεξεργασίας πραγματικού χρόνου. Αναλύσεις αναφορικά με το χρόνο εκτέλεσης δείχνουν ότι η κλασματική εκτίμηση κίνησης αντιστοιχεί στο ήμισυ του συνολικού χρόνου ΕΚ, ο οποίος με τη σειρά του αντιστοιχεί στο 60-90% του συνολικού χρόνου κωδικοποίησης, ανάλογα πάντα με τις επιλεγείσες σχεδιαστικές παραμέτρους [37].

Η κλασματική ΕΚ αυξάνει την ανάλυση μιας τυχαίας περιοχής εντός του πλαισίου της εικονορροής, γεννώντας - με τη μέθοδο της παρεμβολής - νέα δείγματα σε θέσεις ευρισκόμενες μεταξύ των αρχικών εικονοστοιχείων του πλαισίου. Τα σύγχρονα πρότυπα κωδικοαποκωδίκευσης περιγράφουν επακριβώς το φίλτρο παρεμβολής που χρησιμοποιείται από τη μονάδα αντιστάθμισης κίνησης, επιτρέποντας κατ' αυτόν τον τρόπο σε κωδικοποιητή και αποκωδικοποιητή τη δημιουργία και τη χρήση ταυτόσημων πλαισίων αναφοράς. Πιο συγκεκριμένα, το πρότυπο H.264/AVC υπολογίζει τη φωτεινότητα σε ενδιάμεσες θέσεις (κλασματική ακρίβεια $\frac{1}{2}$ του εικονοστοιχείου) μέσω ενός φίλτρου 6 δειγμάτων και μεταξύ των ενδιάμεσων θέσεων (κλασματική ακρίβεια $\frac{1}{4}$ του εικονοστοιχείου) μέσω ενός χαμηλού κόστους φίλτρου 2 δειγμάτων [3]. Παρά το γεγονός ότι η ενσωμάτωση του προτύπου φίλτρου 6 δειγμάτων και στη μονάδα ΕΚ (πριν την αντιστάθμιση) συνιστά κοινή πρακτική των σχεδιαστών κωδικοποιητών, η επιλογή της τεχνικής παρεμβολής για την εύρεση των ανυσμάτων κίνησης (και όχι τον υπολογισμό του υπολοίπου) έγκειται στη διακριτική ευχέρεια του σχεδιαστή αναλόγως των απαιτούμενων επιδόσεων. Κατά συνέπεια είναι δυνατή η χρήση τεχνικών παρεμβολής χαμηλότερης πολυπλοκότητας με στόχο την επιτάχυνση της ΕΚ.

1.2 Συμπύεση εικονορροών με αποδεκίατιση

Τα συμβατικά πρότυπα κωδικοαποκωδίκευσης επιτυγχάνουν εξαιρετικές επιδόσεις σε πληθώρα εφαρμογών για τις οποίες η κωδικοποίηση εκτελείται μόνο μία φορά, από μια ισχυρή μονάδα επεξεργασίας η οποία λειτουργεί χωρίς χρονικούς ή ενεργειακούς περιορισμούς. Από την άλλη πλευρά, η αποκωδικοποίηση είναι σχεδιασμένη να εκτελείται πολλαπλές φορές από ευρέως διαθέσιμες ηλεκτρονικές συσκευές χαμηλού κόστους (π.χ. συσκευές αναπαραγωγής DVD, ψηφιακούς τηλεοπτικούς δέκτες, συσκευές αναπαραγωγής διαδικτυακών πολυμέσων, κ.ά.).

Στις ημέρες μας επίσης αυξάνεται διαρκώς ο αριθμός των εφαρμογών οι οποίες απαιτούν κωδικοποίηση με το ελάχιστο δυνατό κόστος σε χρόνο και ενέργεια, στοχεύοντας στην εξοικονόμηση φορτίου στις μπαταρίες των συσκευών ή στην επίτευξη λειτουργίας σε πραγματικό χρόνο με εξαρτήματα χαμηλού κόστους. Σχετικά παραδείγματα μπορεί κάποιος να συναντήσει σε εφαρμογές οι οποίες χρησιμοποιούν ασύρματες κάμερες χαμηλής κατανάλωσης για παρακολούθηση (π.χ. κυκλοφορίας οχημάτων, άγριων ζώων, ή δασικών εκτάσεων), τηλεδιασκέψεις, ή αποστολή εικονορροών μεταξύ φορητών συσκευών. Προκειμένου να επιτευχθούν όλοι αυτοί οι στόχοι, η έρευνα στράφηκε προς κωδικοποιητές χαμηλής πολυπλοκότητας και σχήματα κωδικοποίησης “αντίστροφης πολυπλοκότητας”, στα οποία ένα τμήμα της πολυπλοκότητας του κωδικοποιητή “μεταφέρεται” στην πλευρά του αποκωδικοποιητή. Χαρακτηριστικό παράδειγμα προς αυτή την κατεύθυνση αποτελεί η κατανεμημένη κωδικοποίηση εικονορροών (Distributed Video Coding [68]).

Έχοντας ως κίνητρο αυτού του είδους τις προσεγγίσεις, η παρούσα διατριβή βασίζεται στην ιδέα της κωδικοποίησης εικονορροών χαμηλότερης ανάλυσης μέσω ενός κωδικοαποκωδικοποιητή στον οποίο τα πλαίσια αρχικά αποδεκατίζονται, στη συνέχεια κωδικοποιούνται, το κωδικοποιημένο δυφίορρευμα (bitstream) αποκωδικοποιείται και τέλος, τα αποκωδικοποιημένα πλαίσια υπερ-δειγματοληπτούνται στο αρχικό τους μέγεθος. Μέσω αυτής της προσέγγισης ο κωδικοποιητής επεξεργάζεται πλαίσια μικρότερων διαστάσεων, με συνέπεια την ελάττωση του χρόνου κωδικοποίησης και του απαιτούμενου επιπέδου κβάντισης για την παραγωγή ενός δυφιορρέυματος συγκεκριμένου όγκου. Η ελάττωση του θορύβου κβάντισης αντισταθμίζει την αρχική απώλεια πληροφορίας, βελτιώνοντας κατ’ αυτόν τον τρόπο την ποιότητα της εικόνας του αποτελέσματος εν συγκρίσει με το υψηλότερα κβαντισμένο δυφίορρευμα των συμβατικών σχημάτων κωδικοποίησης. Η ιδέα του συνδυασμού αποδεκάτισης και κβάντισης έχει ήδη αποδειχθεί πολλά υποσχόμενη για τη συμπίεση εικονορροών χαμηλού δυφιορρυθμού [99][13][92]. Εν συνεχεία, η υπερ-δειγματοληψία στις αρχικές διαστάσεις στην πλευρά του αποκωδικοποιητή μπορεί να υλοποιηθεί μέσω κάποιας τεχνικής από το ερευνητικό πεδίο της Υπερ-Ανάλυσης (ΥΑ, Super-Resolution - SR) [67]. Χαρακτηριστικό των αλγορίθμων ΥΑ αποτελεί η ενίσχυση της ποιότητας της εικόνας εις βάρος του υπολογιστικού φόρτου, με αποτέλεσμα εξαιρετικά σημαντικές διαφορές σε σύγκριση με τις απλές τεχνικές υπερ-δειγματοληψίας. Οι διαφορές αυτές κυμαίνονται από 0.01 και μπορούν να υπερβούν τα 10 dB PSNR σε ασυμπίεστες εικόνες. Επομένως, η χρήση ενός αλγορίθμου ΥΑ στο προτεινόμενο σχήμα θα καθορίσει τόσο την αποδοτικότητα, όσο και την πολυπλοκότητα του κωδικοαποκωδικοποιητή.

1.3 Υπερ-Ανάλυση

Πλέον, όλες οι εφαρμογές οι οποίες σχετίζονται με την επεξεργασία και την ανάλυση εικόνων και εικονορροών, απαιτούν τη δειγματοληψία εικόνων Υψηλής Ευκρίνειας (ΥΕ, High-Resolution - HR). Η δειγματοληψία εικόνων ΥΕ βασίζεται συνήθως στην αύξηση της πυκνότητας και στην ελάττωση μεγέθους των

αισθητήριων διατάξεων δειγματοληψίας. Αμφότερες οι ανωτέρω μέθοδοι εγείρουν σημαντικές απαιτήσεις σε υλικό, αυξάνοντας το κόστος παραγωγής και συντήρησης των αισθητήρων και των φακών εστίασης. Ως λύση στο πρόβλημα αυτό έχει προταθεί η ΥΑ, με την οποία η επιθυμητή εικόνα ΥΕ ανασυντίθεται από μία ή περισσότερες εικόνες Χαμηλής Ευκρίνειας (ΧΕ, Low-Resolution - LR). Προς την κατεύθυνση αυτή, οι λύσεις στη σχετική βιβλιογραφία μοντελοποιούν αρχικά α) το θόρυβο τον οποίο εισαγάγουν οι αισθητήρες/φακοί και β) την κίνηση κατά τη λήψη των εικόνων. Στη συνέχεια, το αποτέλεσμα των ανωτέρω αντισταθμίζεται κατά τη διαδικασία ανακατασκευής [26].

Το πρόβλημα της ΥΑ θεωρείται ένα κακώς τοποθετημένο πρόβλημα, βάσει του ορισμού του Hadamard περί καλώς και κακώς τοποθετημένων προβλημάτων. Αυτό συμβαίνει επειδή δεν υφίσταται μοναδική λύση και επιπλέον, στις περισσότερες των περιπτώσεων κάθε τροποποίηση των αρχικών συνθηκών γεννά και ένα διαφορετικό σύνολο πιθανών λύσεων. Κάθε εικόνα ΧΕ αποτελεί μία παρατήρηση της επιθυμητής λύσης ΥΕ. Στην παρατήρηση αυτή έχει προστεθεί θόρυβος εξαιτίας της εφαρμογής μίας συνάρτησης διασποράς σημείου (ΣΔΣ, Point-Spread Function - PSF) και κατόπιν της αποδεκάτισης του αποτελέσματος. Επιπλέον, κάθε εικόνα ΧΕ ενδέχεται να εμπεριέχει θόρυβο εξαιτίας της κίνησης κατά τη διαδικασία δειγματοληψίας ή ακόμη και της φύσης των απεικονιζόμενων αντικειμένων. Στην περίπτωση που είναι διαθέσιμες πολλαπλές εικόνες ΧΕ, το πρόβλημα κακής τοποθέτησης της ΥΑ βελτιώνεται καθώς περιορίζεται ο χώρος λύσεων. Από την άλλη, αυξάνεται η πολυπλοκότητα και συνεπώς ο χρόνος επεξεργασίας που θα οδηγήσει στην εικόνα ΥΕ. Είναι προφανές επομένως το γεγονός ότι ποιότητα και πολυπλοκότητα συνιστούν αλληλοαναιρούμενους στόχους στο πεδίο της ΥΑ.

1.4 Στόχοι και διάρθρωση της διατριβής

Η τρέχουσα διατριβή μελετά το πεδίο της Υπερ-Ανάλυσης υπό το πρίσμα της μειωμένης πολυπλοκότητας, με σκοπό την παρουσίαση, αποτίμηση επιδόσεων και την υλοποίηση παραλληλοποιησμών μεθόδων και τεχνικών. Επιπλέον, στοχεύει στην εφαρμογή αυτών των μεθόδων σε ένα γενικότερο πλαίσιο επεξεργασίας α) εικόνων και β) εικονορροών με συμπίεση αποδεκάτισης-χβάντισης. Τέλος, μελετά τα απτά αποτελέσματα των τεχνικών αυτών σε πλατφόρμες επιτάχυνσης υλικού.

Στο κεφάλαιο 2 γίνεται αναφορά στη σχετική βιβλιογραφία στο πεδίο της Υπερ-Ανάλυσης, της συμπίεσης εικονορροών με αποδεκάτιση και χβάντιση και την επιτάχυνση αλγορίθμων σε Μονάδες Επεξεργασίας Γραφικών (MEΓ, Graphics Processing Units - GPUs) και Συστοιχίες Επιτόπια Προγραμματιζόμενων Πυλών (Field-Programmable Gate Arrays - FPGAs). Στο κεφάλαιο 3, παρουσιάζονται τεχνικές χαμηλής πολυπλοκότητας για την υπερ-δειγματοληψία εικόνας, με στόχο την εφαρμογή τους στην εκτίμηση κίνησης του προτύπου H.264/AVC [3]. Στο κεφάλαιο 4, παρουσιάζεται η βελτίωση και εφαρμογή των τεχνικών αυτών στο πλαίσιο της Υπερ-Ανάλυσης από μία εικόνα. Κατόπιν, στο Κεφάλαιο 5 παρουσιάζουμε τον αλγόριθμο *L-SEABI*, έναν ολοκληρωμένο αλγόριθμο ΥΑ

από μία εικόνα, ο οποίος αποτιμάται ως αυτόνομη λύση αλλά και στο πλαίσιο της συμπίεσης εικονορροών με αποδεκτικότητα. Πρωτεύων στόχος είναι η βελτίωση της απόδοσης του διπόλου δυοφωριθμίου-παραμόρφωσης του κωδικοποιητή και, παράλληλα, η ελάττωση της συνολικής πολυπλοκότητας του συστήματος.

Εκτός από απόδοση και πολυπλοκότητα, εξετάζουμε και παραμέτρους όπως η φορητότητα και η προσαρμοστικότητα, μέσω των οποίων επιτρέπεται η ταχεία ανάπτυξη/χρήση του συστήματος σε πληθώρα πραγματικών σεναρίων. Για το σκοπό αυτό, υιοθετούμε το σχήμα αποδεκτικότητας-κβάντισης με διακριτές μονάδες μετασχηματισμού διαστάσεων και κωδικοποίησης. Αυτό επιτρέπει την εύκολη τροποποίηση του συστήματος, καθώς οποιαδήποτε δομική μονάδα μπορεί να αντικατασταθεί από μονάδες προσαρμοσμένες στις απαιτήσεις της εκάστοτε εφαρμογής. Πληθώρα τέτοιων μονάδων είναι ήδη διαθέσιμη από τη βιβλιογραφία και τη βιομηχανία, τόσο για το μετασχηματισμό μεγέθους, όσο και για την υπερδειγματοληψία. Επιπλέον, η απόφασή μας ελήφθη βάσει ενός σημαντικού συμπεράσματος από την τρέχουσα βιβλιογραφία: ότι στις περισσότερες περιπτώσεις, ο συνδυασμός της αποδεκτικότητας με άλλες βασικές διεργασίες εντός του κωδικοποιητή αυξάνει την πολυπλοκότητα του συστήματος και, συνεπώς, βρίσκεται εκτός των στόχων της παρούσας διατριβής.

Η παρούσα διατριβή εστίασε στην ελάττωση της πολυπλοκότητας του συστήματος συνδυάζοντας δύο προσεγγίσεις. Πρώτον, θεωρούμε ότι ο κωδικοποιητής επεξεργάζεται πλαίσια χαμηλότερης ανάλυσης: η αποδεκτικότητα κατά παράγοντα δύο σε κάθε διάσταση (η βέλτιστη τιμή βάσει της βιβλιογραφίας) σχεδόν υποτετραπλασιάζει τον υπολογιστικό φόρτο του κωδικοποιητή. Δεδομένου ότι ο τελευταίος αποτελεί την απαιτητικότερη υπολογιστικά μονάδα του συστήματος, αυτό εν δυνάμει υποτετραπλασιάζει την πολυπλοκότητα του συνολικού συστήματος. Αυτό αποτελεί σημαντικό κέρδος ειδικά στην επεξεργασία εικονορροών υψηλής ή εξαιρετικά υψηλής ευκρίνειας (High Definition - HD, Ultra High Definition - UHD). Δεύτερον, εστιάζουμε προσεκτικά στην επιλογή του αλγορίθμου υπερδειγματοληψίας στην πλευρά του αποκωδικοποιητή. Διερευνούμε το πεδίο ΥΑ αναπτύσσοντας και συγκρίνοντας μεταξύ τους διάφορες αντιπροσωπευτικές τεχνικές, ως προς τις επιδόσεις ποιότητας και το χρόνο εκτέλεσής τους. Έπειτα εξετάζουμε το αποτέλεσμά τους στο προτεινόμενο σχήμα συμπίεσης και προτείνουμε τον αλγόριθμο *L-SEABI*, ένα νέο αλγόριθμο ΥΑ ο οποίος βελτιώνει την ποιότητα του σχήματος με μικρό σχετικά υπολογιστικό κόστος.

Στο κεφάλαιο 6 παρουσιάζουμε τεχνικές παραλληλοποίησης και επιτάχυνσης του προτεινόμενου αλγορίθμου ΥΑ σε πλατφόρμες υλικού. Στο ίδιο κεφάλαιο πραγματοποιούμε εξερεύνηση του χώρου σχεδίασης και εκτεταμένη αποτίμηση της απόδοσης του προτεινόμενου αλγορίθμου, εξαγάγοντας χρήσιμα συμπεράσματα και για τη συσχέτιση των πλατφορμών υλοποίησης. Εν συνεχεία, στο Κεφάλαιο 7 εξετάζουμε τη χρήση του προτεινόμενου αλγορίθμου ως μέσου βελτίωσης περισσότερο περίπλοκων αλγορίθμων ΥΑ και τέλος, την εφαρμογή των προτεινόμενων τεχνικών παραλληλοποίησης σε περισσότερο περίπλοκες μεθόδους επεξεργασίας εικόνων. Τέλος, στο κεφάλαιο 8 παρουσιάζουμε μια σύνοψη των συμπερασμάτων της διατριβής και αναφερόμαστε στους μελλοντικούς ερευ-

ΚΕΦΑΛΑΙΟ 1

1.4. Στόχοι και διάρθρωση της διατριβής

νητικών στόχους.

ΚΕΦΑΛΑΙΟ 2

Πρόβλημα και Σχετική Βιβλιογραφία

2.1 Συμπίεση με αποδεκίαση και κβάντιση

Οι απωλεστικές τεχνικές συμπίεσης βασίζονται σε δύο βασικές αρχές: είτε στην απόρριψη τμημάτων του δειγματοληπτημένου σήματος, είτε στην αποδεκίαση της εισόδου με σκοπό την παραγωγή σήματος μικρότερου όγκου πληροφορίας. Στις περισσότερες περιπτώσεις, το επεξεργασμένο σήμα θα παρουσιάζει θόρυβο κβάντισης/αποδεκίασης. Βέβαια, τις τελευταίες δεκαετίες, η έρευνα στο πεδίο της συμπίεσης εικόνων/εικονορροών έχει επιτύχει σημαντικά βήματα προς την ελάττωση του θορύβου κβάντισης με τη βοήθεια υπολογιστικά απαιτητικών προτύπων κωδικοαποκωδίκευσης όπως τα H.264/AVC και HEVC [64][83]. Επίσης, η εμφάνιση και ανάπτυξη περίπλοκων αλγορίθμων αποδεκίασης και υπερδειγματοληψίας εικόνων βοήθησε στην περιστολή του θορύβου αποδεκίασης. Πιο πρόσφατα, η έρευνα τείνει να στρέφεται στο συνδυασμό προηγμένων τεχνικών από τα δύο αυτά πεδία με σκοπό να εξισορροπήσει τους δύο προαναφερθέντες τύπους θορύβου και κατά συνέπεια να παράξει σήμα υψηλότερης ποιότητας για τον ίδιο δυφορρυθμό. Παραδείγματα από τη σχετική βιβλιογραφία έχουν δείξει πως ένας τέτοιος συνδυασμός δύναται όντως να βελτιώσει την αποδοτικότητα των κωδικοποιητών, ειδικά στην περίπτωση συμπίεσης με χαμηλό δυφορρυθμό [99][13][58][53][11] [39][59][100][54][92][81] [8][66][48][21].

Το όφελος της συμπίεσης αποδεκιασμένων εικόνων τονίστηκε από τους Zeng και Venetsanopoulos το 1993 [99] στο πλαίσιο συμπίεσης εικόνων με το πρότυπο JPEG. Παρόμοια προσέγγιση ακολούθησε και η έρευνα μία δεκαετία αργότερα [13][100][53]. Για παράδειγμα, οι συγγραφείς στην [13] παρουσίασαν ένα αναλυτικό μοντέλο με τη βοήθεια του οποίου εξήγησαν το παρατηρηθέν κέρδος της εφαρμογής αποδεκίασης πριν από το μετασχηματισμό και την κωδικοποίηση του σήματος. Τα πειραματικά τους ευρήματα επιβεβαίωσαν τη βελτίωση στην ποιότητα της εικόνας για συγκεκριμένο εύρος δυφορρυθμού, τόσο αντικειμενικά, όσο και υποκειμενικά.

Όσον αφορά τη συμπίεση εικονορροών, οι ερευνητικές ομάδες των Barreto κ.ά. [11] και Pgin κ.ά. [39] αξιοποίησαν τον επανακαθορισμό του μεγέθους των πλαισίων βελτιώνοντας κατ' αυτόν τον τρόπο τις επιδόσεις του διπλού δυφορρυθμού-

παραμόρφωσης. Οι συγγραφείς στην [39] μεταβάλλουν τις διαστάσεις της εικόνας στο χώρο διακριτού μετασχηματισμού συνημιτόνου (ΔΜΣ, Discrete Cosine Transform - DCT), εξωτερικά του κωδικοποιητή H.264. Με άλλα λόγια, η μονάδα πρόβλεψης του κωδικοποιητή επιδρά επί του αρχικού πλαισίου και προωθεί το αποτέλεσμα σε μία μονάδα μετασχηματισμού, η οποία πραγματοποιεί αποδεκίαση απορρίπτοντας τις υψίσυχνες συνιστώσες του ΔΜΣ. Σε σχέση με τη συμβατική κωδικοποίηση μέσω H.264/AVC, το σχήμα των συγγραφέων στο [39] παρουσιάζει βελτίωση ποιότητας κατά 2dB ως προς τη μετρική PSNR για δυοιορρυθμούς κάτω των 20 Kbps. Οι συγγραφείς στην [59] εξελίσσουν περαιτέρω την παραπάνω ιδέα, προτείνοντας την προσαρμοστική αποδεκίαση ανά πλακίδιο εικόνας (Macro-Block - MB) βάσει του περιεχομένου (της μεταβλητότητας των συντελεστών ΔΜΣ). Εν συγκρίσει με το συμβατικό MPEG-2 πρότυπο, τα αποτελέσματα τους παρουσιάζουν βελτίωση 1-2 dB (μετρική PSNR) για δυοιορρυθμούς έως 100 Kbps. Επίσης οι συγγραφείς στην [81] επιλέγουν να μεταβάλλουν τις διαστάσεις μόνο στο χώρο δια-πλαισιακής πρόβλεψης (Inter-Frame Prediction) επιτρέποντας την ανασύσταση από εικόνες χαμηλότερου θορύβου. Εν συγκρίσει με το H.264/AVC, η προταθείσα λύση ενισχύει την ποιότητα κατά μέσο όρο κατά 1.5-2 dB για εικονορροές δυοιορρυθμού 30-60 Kbps. Μειονέκτημα όλων αυτών των τεχνικών που συνδυάζουν αποδεκίαση και μονάδες μετασχηματισμού, αποτελεί η σημαντική αύξηση της πολυπλοκότητας του κωδικοποιητή (ο χρόνος κωδικοποίησης σχεδόν τριπλασιάζεται σε σχέση με τη συμβατική κωδικοποίηση).

Ενδιαφέρον παρουσιάζει η μέθοδος των συγγραφέων στο [11], οι οποίοι προτείνουν τον τεμαχισμό της εικόνας και την προσαρμογή της μεθόδου μετασχηματισμού διαστάσεων ανά πλακίδιο, με βάση τις προβλέψεις του προτύπου MPEG-4 και την ταξινόμηση κίνησης ανά περιοχές εικόνας. Αυτή η μέθοδος παρουσιάζει εφαρμογή όταν δε λαμβάνεται υπόψιν η συνολική πολυπλοκότητα, ενώ επιπλέον απαιτεί και τη μετάδοση πρόσθετης πληροφορίας προς τον αποκωδικοποιητή. Η προσέγγιση αυτή επιτυγχάνει εντυπωσιακή αύξηση στη μετρική PSNR έως και κατά 4 dB, επεκτείνοντας το εύρος ζώνης δυοιορρυθμού στην περιοχή του 1 Mbps. Ωστόσο, αυξάνει σημαντικά τον υπολογιστικό φόρτο (ο χρόνος κωδικοποίησης σχεδόν εξαπλασιάζεται).

Όλες οι προαναφερθείσες τιμές κέρδους ποιότητας αφορούν εικονορροές με μικρές διαστάσεις πλαισίου (π.χ. 352x288 εικονοστοιχεία). Η αύξηση στις διαστάσεις των πλαισίων βελτιώνει την αποδοτικότητα κωδικοποίησης των σχημάτων αποδεκίασης-κβάντισης, κυρίως λόγω του ότι συνεπάγεται αύξηση στη χωρική συσχέτιση μεταξύ των εικονοστοιχείων και κατ' επέκταση ελάττωση του θορύβου αποδεκίασης. Τέτοιου είδους σχήματα αξιοποιήθηκαν από πρόσφατες έρευνες για τη συμπίεση εικονορροών διαστάσεων 1280 × 720 και άνω [58][54][92][88]. Οι συγγραφείς της [54] εξελίσσουν την προσέγγιση του [59] ελέγχοντας τη μέθοδο αποδεκίασης ανά πλακίδιο μέσω τεχνικών αμοιβαίας βελτιστοποίησης του διπλόου δυοιορρυθμού-κβάντισης (rate-distortion-optimization - RDO). Η υπερδειγματοληψία την οποία εφαρμόζουν προσαρμόζεται στη δραστηριότητα κίνησης κάθε πλακιδίου, βάσει των συντελεστών του ΔΜΣ και της τιμής αθροίσματος απολύτων διαφορών (ΑΑΔ, sum of absolute differences - SAD). Ο αποκωδικο-

ποιητής τους επιλέγει μεταξύ τριών διαφορετικών μεθόδων υπερ-δειγματοληψίας: προσαρμοστικής ΥΑ, υπερ-δειγματοληψίας με διατήρηση ακμών, ή δικυβικής παρεμβολής. Τέτοιου επιπέδου προσαρμοστικότητα προφανώς απαιτεί σημαντικές τροποποιήσεις στο πρότυπο κωδικοαποκωδίκευσης, αφού περιλαμβάνει υπολογισμούς και αποφάσεις σε πραγματικό χρόνο, εξαρτώμενες από τα αποτελέσματα του συμβατικού συμπίεστη. Στο [92], παρουσιάστηκε μια διαφορετική προσέγγιση η οποία αποσυζευγνύει τελείως τις διαδικασίες μετασχηματισμού διαστάσεων και κωδικοποίησης. Οι συγγραφείς των [92] και [21] μελετούν προσαρμοστικές μεθόδους αποδεκτίσης στο πεδίο του χώρου προκειμένου να συμπεράνουν τη βέλτιστη τιμή του παράγοντα αποδεκτίσης για δεδομένο δυοφορρυθμό. Οι συγκρίσεις διαφόρων παραγόντων αποδεκτίσης της έρευνας στο [92] δείχνουν ότι σε μη προσαρμοστικά σενάρια, η βέλτιστη απόδοση επιτυγχάνεται με αποδεκτίση κατά 2 σε κάθε διάσταση ($1/2$ της οριζόντιας και $1/2$ της κατακόρυφης διάστασης). Τα αποτελέσματα της έρευνας αυτής σε εικόνες διαστάσεων 1280×720 και άνω με δικυβική παρεμβολή, παρουσιάζουν κέρδος 2-4 dB (PSNR) για εικονορροές με δυοφορρυθμό κάτω των 25 Mbps. Οι συγγραφείς της [88] τροποποιούν σε βάθος τον κωδικοποιητή, έτσι ώστε να επεξεργάζεται κάθε πλακίδιο σε υψηλή και χαμηλή ευκρίνεια, επιλέγοντας τελικά την κωδικοποίηση βάσει συγκρίσεων στο προβλεπόμενο κόστος δυοφορρυθμού-παραμόρφωσης. Όπως είναι αναμενόμενο, αυτή η προσέγγιση επεκτείνει σημαντικά το ωφέλιμο εύρος δυοφορρυθμού του σχήματος σε υψηλές τιμές (όπου και ο ευέλκτος κωδικοποιητής των συγγραφέων εκφυλίζεται στο συμβατικό πρότυπο H.264/AVC), αυξάνοντας προφανώς την υπολογιστική πολυπλοκότητα. Τέλος, να σημειώσουμε ότι η κωδικοποίηση με αποδεκτίση και υπερ-δειγματοληψία έχει χρησιμοποιηθεί και σε επεκτάσεις προτύπων συμπίεσης όπως το Scalable Video Coding (SVC) και το Multi-view Video plus Depth (MVD). Στην περίπτωση του SVC, ο θόρυβος αποδεκτίσης αντισταθμίζεται με την κωδικοποίηση απομένουσας πληροφορίας, ενώ στο MVD συνήθως κωδικοποιείται μία αποδεκατισμένη εκδοχή του χάρτη βάθους κάθε πλαισίου [80].

2.2 Αλγόριθμοι Υπερ-Ανάλυσης

Η αποδοτικότητα της κωδικοποίησης σε όλα τα σχήματα συμπίεσης με αποδεκτίση και κβάντιση βασίζεται με τη σειρά της στην αποδοτικότητα των αλγορίθμων που χρησιμοποιούνται για το μετασχηματισμό των διαστάσεων των πλαισίων. Βασική αρχή των αλγορίθμων ΥΑ είναι η ανακατασκευή της επιθυμητής εικόνας υψηλής ευκρίνειας (ΥΕ) από μία - ή περισσότερες - εικόνες/παρατηρήσεις χαμηλής ευκρίνειας (ΧΕ). Η διαδικασία μπορεί να πραγματοποιηθεί απευθείας στο επίπεδο των εικονοστοιχείων (χωρικό πεδίο) ή μέσω των συντελεστών μετασχηματισμού (συχνοτικό πεδίο). Κατά την τελευταία δεκαετία, η ερευνητική κοινότητα εστίασε στην αποτελεσματική αντιμετώπιση δύο βασικών προβλημάτων: πρώτον, τη μειωμένη ανθεκτικότητα σε σφάλματα λόγω ανακριβειών μοντελοποίησης και δεύτερον, το διαρκώς διογκούμενο υπολογιστικό κόστος.

Η ακριβής μοντελοποίηση του θορύβου των αισθητήρων δειγματοληψίας και

των φακών καθώς και της υφιστάμενης κίνησης κατά τη λήψη της εικόνας είναι απαραίτητη για την αντιστάθμιση όλων των παραπάνω κατά τη διαδικασία ανακατασκευής [26]. Η πλειονότητα των τεχνικών ΥΑ βασίζεται στο μοντέλο υποβάθμισης εικόνων [26]. Κατά κύριο λόγο, οι τεχνικές ΥΑ πραγματοποιούν καταχώρηση εικόνων (αντιστοίχιση εικόνων μέσω μετασχηματισμών), παρεμβολή και τέλος αφαίρεση θορύβου/θόλωσης προκειμένου να παράξουν την εικόνα ΥΕ [27], [72], [30]. Οι απλές μέθοδοι ΥΑ πραγματοποιούν παρεμβολή στο χωρικό πεδίο μέσω ενός εκ των γνωστών τεχνικών, όπως η *διγραμμική*, η *δικυβική*, ή η *Lanczos* [46] [14].

Για τη βελτίωση του αποτελέσματος ΥΑ απαιτείται η ύπαρξη ενός μοντέλου το οποίο λαμβάνει υπόψιν του τις εξαρτήσεις μεταξύ της καταχώρησης, της παρεμβολής και της αφαίρεσης του θορύβου. Οι Irani και Peleg πρότειναν το σχήμα της *επαναληπτικής οπισθο-προβολής (EOΠ, Iterative Back-Projection - IBP)* [40], το οποίο για να παράξει την εικόνα ΥΕ βασίζεται στο μοντέλο υποβάθμισης και οπισθο-προβάλλει συνθηκοπαγώς έναν πίνακα σφάλματος στην εικόνα. Το γενικευμένο μοντέλο *EOΠ* μπορεί να περιλαμβάνει πληροφορία από πολλαπλές εικόνες ΧΕ οι οποίες έχουν παραχθεί από στοχαστικά μοντέλα στα οποία γίνεται εκτίμηση κίνησης, στροφής, μετατόπισης και θορύβου. Στην περίπτωση που οι εικόνες ΧΕ δεν είναι διαθέσιμες, μπορούν να προσομοιωθούν, ή - στην περίπτωση εικονορροών - να θεωρηθούν ως διαδοχικά πλαίσια στα οποία περιέχεται πληροφορία κίνησης [73]. Σε γενικές γραμμές, το σχήμα *EOΠ* βελτιώνει την ποιότητα των απλών μεθόδων παρεμβολής κατά 1-5 dB (PSNR) [22], απαιτώντας μία έως τρεις τάξεις μεγέθους περισσότερο χρόνο εκτέλεσης.

Οι υπερσύγχρονες τεχνικές ΥΑ συνδυάζουν στοχαστικά μοντέλα με αραιή αναπαράσταση πρότερης γνώσης προκειμένου να εξαγάγουν μη διαθέσιμη πληροφορία και να αντιμετωπίσουν τον κακώς-ορισμένο χαρακτήρα της ΥΑ. Για παράδειγμα, η ερευνητική ομάδα των Katsaggelos κ.ά. πρότεινε ένα Bayesian πλαίσιο για τη μοντελοποίηση των συσχετίσεων μεταξύ της εικόνας ΥΕ και πολλαπλών παρατηρήσεων ΧΕ. Αρχικά, πρότειναν τη χρήση πρότερης γνώσης με νόρμα l_1 [90] και επέκτειναν το μοντέλο τους για το συνδυασμό με την Ταυτόχρονη Αυτο-Παλινδρομική (Simultaneous Auto Regressive - SAR) μη αραιή πρότερη γνώση [89]. Πρόσφατα ανέπτυξαν ένα μεταβολικό Bayesian πλαίσιο με βελτιωμένη ανθεκτικότητα σε σφάλματα το οποίο υπολογίζει προσαρμοστικά τις παραμέτρους υποβάθμισης της εικόνας [10]. Αυτές οι μέθοδοι παρέχουν την υψηλότερη ποιότητα ΥΑ, ωστόσο το υπολογιστικό τους κόστος αυξάνεται σημαντικά με την αύξηση των παρατηρήσεων ΧΕ και του παράγοντα υπερ-δειγματοληψίας. Το κόστος είναι ακόμη υψηλότερο όταν οι παράμετροι υποβάθμισης δεν έχουν προϋπολογιστεί και πρέπει να υπολογιστούν κατά την εκτέλεση.

Σε διαφορετική κατεύθυνση, οι πιο σύγχρονες μέθοδοι συνδυάζουν ΥΑ αραιής αναπαράστασης με λεξικά. Η ομάδα των Ma κ.ά. χρησιμοποιούν αραιή αναπαράσταση για τη συσχέτιση του λεξικού ΧΕ με τα πλακίδια ΥΕ [98]. Τη μέθοδό τους επεκτείνουν οι Wu κ.ά. με α) προσαρμοστική επιλογή του χώρου αραιής αναπαράστασης και β) κανονικοποίηση για την επίτευξη σύγκλισης [23]. Αυτού του τύπου οι προσεγγίσεις είναι εξαιρετικά αποτελεσματικές, ακόμη και με μία

παρατήρηση ΧΕ. Από την άλλη πλευρά, το υπολογιστικό τους κόστος είναι εξαιρετικά υψηλό (6 τάξεις μεγέθους υψηλότερος χρόνος εκτέλεσης σε σχέση με απλές μεθόδους παρεμβολής). Συν τοις άλλοις, η χρήση λεξικών πλακιδίων απαιτεί τη χρήση παράπλευρης πληροφορίας κατά την υπερ-δειγματοληψία, η οποία στην περίπτωση του σχήματος αποδεκτικής-κβάντισης πρέπει να μεταδοθεί από την πλευρά του κωδικοποιητή - δεσμεύοντας έτσι ένα τμήμα του εύρους ζώνης της εφαρμογής.

2.3 Επιτάχυνση αλγορίθμων σε υλικό

2.3.1 Επιτάχυνση σε Μονάδες Επεξεργασίας Γραφικών

Η βιβλιογραφία προσφέρει πληθώρα παραδειγμάτων στα οποία γίνεται χρήση πολυπύρηνων αρχιτεκτονικών ως αποδοτικός τρόπος για την επιτάχυνση της επεξεργασίας εικόνων και εικονορροών. Στο πεδίο της ΥΑ, η ομάδα των Freedman κ.ά. πρότεινε και εφάρμοσε μία μέθοδο ΥΑ από από μία παρατήρηση στη ΜΕΓ GTX 480 προσεγγίζοντας υπερ-δειγματοληψία σε πραγματικό χρόνο κατά το μετασχηματισμό εικόνων από 640×360 σε 1920×1080 . Οι συγγραφείς της λεξικοπαγούς μεθόδου στη [101] χρησιμοποιούν επίσης ΜΕΓ για την επιτάχυνση της μεθόδου τους. Σε μία διαφορετική κατεύθυνση, οι συγγραφείς των [19] και [20] αξιοποίησαν τη βιβλιοθήκη *cuda-convnet* [6] για την εκπαίδευση του Συνελικτικού Νευρωνικού Δικτύου τους μέσω του οποίου αντιστοιχίζουν εικόνες ΧΕ και ΥΕ. Πιο πρόσφατα, οι Jung κ.ά. χρησιμοποίησαν τη βιβλιοθήκη *cuFFT* [62] στις αποσυνελικτικές μεθόδους ΥΑ τους, επιτυγχάνοντας υπερ-δειγματοληψία κατά δύο σε ΥΕ 1920×1080 εντός 53 ms με τη βοήθεια μιας ΜΕΓ GTX 580. Η χρήση ΜΕΓ αποτέλεσε εξάλλου από νωρίς κοινό τόπο στην επιτάχυνση εφαρμογών στερεοσκοπικής όρασης [28]. Το box φίλτρο στην ανωτέρω έρευνα επιτυγχάνει επιδόσεις πραγματικού χρόνου κατά την επεξεργασία εικόνων ανάλυσης 1024×768 με 50 επίπεδα ανομοιότητας στη ΜΕΓ 8800GTX. Στην ίδια κατηγορία εφαρμογών, οι συγγραφείς των [49] [50] συνδυάζουν χωρική και χρονική επεξεργασία, επιτυγχάνοντας επεξεργασία πραγματικού χρόνου σε εικόνες ανάλυσης 320×240 με πιο εξελιγμένες ΜΕΓ.

2.3.2 Επιτάχυνση με Συστοιχίες Επιτόπια Προγραμματιζόμενων Πυλών (ΣΕΠΠ, Field-Programmable Gate Arrays - FPGAs)

Η αντιστοίχιση των αλγοριθμικών διαδικασιών σε ολοκληρωμένα κυκλώματα (OK, Integrated Circuits - ICs) φαίνεται ως η πιο προφανής λύση στις περιπτώσεις κατά τις οποίες τα υπολογιστικά χαρακτηριστικά της εφαρμογής απαιτούν το συνδυασμό υψηλών επιδόσεων με χαμηλή ενεργειακή κατανάλωση. Οι συστοιχίες επιτόπια προγραμματιζόμενων πυλών αποτελούνται από τροποποιήσιμες δομικές βαθμίδες λογικών πράξεων καθώς και κυκλώματα ενσωματωμένων

συνδυαστικών λειτουργιών τα οποία καλύπτουν τέτοιου είδους υπολογιστικές απαιτήσεις με εξαιρετική ευελιξία, ειδικά όσον αφορά εφαρμογές υπολογιστικής όρασης. Η ομάδα των Bowen κ.ά. για παράδειγμα [12], συνδύασε μία βαθμίδα σταθμισμένης μέσης τιμής με επαναληπτική βελτίωση προκειμένου να επιτύχει υπερ-δειγματοληψία από πολλαπλά πλαίσια XE σε ανάλυση 720p με ρυθμό 61 πλαισίων ανά δευτερόλεπτο (π.α.δ., frames per second - fps) στη συστοιχία FPGA Xilinx Virtex II. Οι συγγραφείς της [7] υλοποιούν ένα σχήμα EOP στη συστοιχία Virtex 4 επιτυγχάνοντας απόδοση 25 π.α.δ. με δύο επαναλήψεις και αποτέλεσμα 1024×1024 εικονοστοιχείων. Προσφάτως, η ομάδα των Sanada κ.ά. [74] πρότεινε μία αρχιτεκτονική ΥΑ βασισμένη σε πράξεις ακραίων με ακμοενισχυτικές τεχνικές και απόδοση 60 π.α.δ. ανάλυσης 400×400 εικονοστοιχείων σε συστοιχίες Altera ep2s130 Stratix II. Η ομάδα των Pérez κ.ά. σχεδίασε μία αρχιτεκτονική επεξεργασίας ροής δεδομένων (stream-processing) για την ΥΑ δεδομένων από συστοιχίες μικροφακών σε κάμερες πεδίου φωτός [69]. Η προτεινόμενη λύση απαιτεί 105.9 ms για να παράξει εικόνες 589×589 από μικροφακούς 291×291 σε πλατφόρμες FPGA χαμηλής κατανάλωσης. Οι συγγραφείς της [65] επεκτείνουν το πρόβλημα σε αναλύσεις υψηλής ευκρίνειας, προτείνοντας μία λύση ΥΑ πραγματικού χρόνου σε συστοιχία Altera Aria II. Σε απαιτητικότερα προβλήματα υπολογιστικής όρασης, η ομάδα των Greisen κ.ά. προτείνει υλοποίηση στερεοσκοπικής αντιστοίχισης σε FPGA, με την οποία πραγματοποιείται επεξεργασία 256 επιπέδων ανομοιότητας ανάλυσης 1080p σε πραγματικό χρόνο [32]. Για να το επιτύχουν, βασίζονται σε μία ιεραρχική προσέγγιση η οποία αρχικά υπολογίζει την ανομοιότητα για εικόνες 16 φορές μικρότερες και εν συνεχεία βελτιώνει το χάρτη βάθους με τοπικές αναζητήσεις ανά περιοχές εικονοστοιχείων. Γενικό χαρακτηριστικό των υλοποιήσεων στερεοσκοπικής όρασης πραγματικού χρόνου είναι το αυξημένο κόστος σε πόρους υλικού [42, 94].

2.3.3 Συγκριτικές μελέτες επιτάχυνσης σε Μονάδες Επεξεργασίας Γραφικών και Συστοιχίες Επιτόπια Προγραμματιζόμενων Πυλών

Σε όλη τη βιβλιογραφία, η συγκριτική αποτίμηση των επιδόσεων μεταξύ MEG και FPGA βασίζεται στην επιτάχυνση γνωστών αλγορίθμων, ή την υλοποίηση με γενικευμένες μεθόδους κάποιας συγκεκριμένης κατηγορίας αλγορίθμων. Η ομάδα των Che κ.ά. για παράδειγμα, [15] χρησιμοποιούν μία MEG 8800GTX και μία συστοιχία Virtex II για να εξετάσουν της επιδόσεις τεχνικών όπως η Γκαουσιανή Εξάλειψη (Gaussian Elimination) από τη γραμμική άλγεβρα, το κρυπτογραφικό πρότυπο DES και ο αλγόριθμος Needleman-Wunsch για την ευθυγράμμιση γενετικών ακολουθιών. Στη δημοσίευση [97] προτείνεται μία πλήρως τροποποιημένη αποδόμηση Cholesky στη MEG GTX 480 και τη συστοιχία πυλών Virtex 6 xc6vnx475t. Οι συγγραφείς της [43] εξετάζουν ένα ευρύ φάσμα εφαρμογών (γεννήτριες ψευδοτυχαίων αριθμών, πολλαπλασιασμό πινάκων, παράλληλη περιστολή και εξομοίωση N σωματιδίων) σε μία MEG GTX 285 και την πλατφόρμα πολυσυστοιχιών FPGA HC-1. Στο πεδίο της υπολογιστικής όρασης, οι Kalarot κ.ά. [44] υλοποιούν ένα γενικό αλγόριθμο ανομοιότητας στη MEG GTX 280 και τη

συστοιχία Altera Stratix III. Στη δημοσίευση [76] παρουσιάζεται ένα πλήρες σύστημα ανίχνευσης αιμοφόρων αγγείων από ιατρικές εικόνες στη MEG GTX 295 και τη συστοιχία Spartan-3. Σε ακόμη πιο πρόσφατες εξελίξεις, οι Pietron κ.ά. [70] συγκρίνουν ένα σύστημα ταξινόμησης ανθρώπινου δέρματος υλοποιημένο στη MEG Tesla m2090 και σε συστοιχίες Virtex 5, ενώ η [56] αποτιμά την απόδοση ενός αλγορίθμου εντοπισμού ελαττωμάτων κεραμικών πλακιδίων στη MEG 9800GT και τρεις διαφορετικές κατηγορίες FPGA. Σε διαφορετική κατεύθυνση, οι συγγραφείς της [34] αποτιμούν την απόδοση κώδικα στερεοσκοπικής αντιστοίχισης μέσω αυτοματοποιημένης σύνθεσης υψηλού επιπέδου από MEG σε FPGA.

Η ανωτέρω διερεύνηση της σχετικής βιβλιογραφίας οδηγεί στο συμπέρασμα (το οποίο επιβεβαιώνεται και από την παρούσα διατριβή) ότι οι MEG είναι καταλληλότερες για υπολογισμούς Μίας Εντολής, Πολλαπλών Δεδομένων (MEΠΔ, Single Instruction Multiple Data - SIMD) όταν δεν υφίστανται αλγοριθμικές ενδοεξαρτήσεις. Η ανάπτυξη σε MEG επιτρέπει ευκολότερη μετάβαση από συμβατικές υλοποιήσεις σε λογισμικό, παρέχει ευελιξία τροποποιήσεων στο σχεδιαστή και χαρακτηρίζεται από χαμηλό κόστος, τόσο για την απόκτηση, όσο και για τη συντήρηση. Από την άλλη πλευρά, οι MEG είναι ενεργοβόρες και λιγότερο αποδοτικές όταν η εφαρμογή απαιτεί α) μεγάλο πλήθος / ακανόνιστες προσπελάσεις στη μνήμη, β) εκτεταμένα στάδια συγχρονισμού. Οι συστοιχίες πυλών είναι καταλληλότερες για εφαρμογές επεξεργασίας ροών δεδομένων, παρέχοντας σαφώς υψηλότερο επίπεδο ελέγχου στις λεπτομέρειες της υλοποίησης. Στα πλεονεκτήματά τους συγκαταλέγεται ο συνδυασμός υψηλής ρυθμοαπόδοσης, χαμηλής κατανάλωσης με αιτιοκρατικές επιδόσεις και κόστος σε πόρους υλοποίησης. Ωστόσο ο προγραμματισμός τους είναι περισσότερο περίπλοκος, απαιτούν ενδελεχή κατανόηση του υπό υλοποίηση αλγορίθμου και είναι ακατάλληλες όταν η εφαρμογή διακρίνεται από α) περίπλοκη ροή δεδομένων ή β) πολλά δεδομένα προς επεξεργασία εξαιτίας των περιορισμένων πόρων μνήμης τους.

ΚΕΦΑΛΑΙΟ 3

Τεχνικές Φίλτρων Παρεμβολής για Εκτίμηση Κίνησης με Εφαρμογή σε Κωδικοποιητές H.264/AVC

Η παρούσα ενότητα δομείται επάνω στη ιδέα της υλοποίησης τεχνικών παρεμβολής δυνητικά χαμηλότερης πολυπλοκότητας αντί του πρότυπου φίλτρου 6 δειγμάτων στον κωδικοποιητή H.264/AVC. Ανώτερος στόχος μας είναι η επιτάχυνση της διαδικασίας Εκτίμησης Κίνησης (ΕΚ) η οποία χαρακτηρίζεται από σημαντικά υψηλότερο υπολογιστικό φόρτο εν συγκρίσει με την Αντιστάθμιση Κίνησης (ΑΚ). Για το σκοπό αυτό υλοποιούμε διάφορες τεχνικές παρεμβολής τις οποίες και ενοποιούμε με τη μονάδα ΕΚ. Παράλληλα, διατηρούμε την πρότυπη ΑΚ του H.264/AVC προκειμένου να μετρήσουμε τον αντίκτυπο των τεχνικών αυτών στο αποτέλεσμα της μονάδας πρόβλεψης.

Στη σχετική βιβλιογραφία περιλαμβάνονται τόσο προτάσεις οι οποίες αποφεύγουν/αντικαθιστούν τους υπολογισμούς του προτύπου [33][38], όσο και λύσεις με τη μορφή αρχιτεκτονικών οι οποίες στοχεύουν στην αποδοτική υλοποίηση της διαδικασίας φιλτραρίσματος του H.264 και την κάλυψη των αυξημένων απαιτήσεων της σε μνήμη [16] [96] [45]. Σε αυτό το σημείο να σημειώσουμε ότι το πρότυπο H.264 ορίζει τον πυρήνα $\langle 1, -5, 20, 20, -5, 1 \rangle$, ο οποίος πολλαπλασιάζεται με 6 διαδοχικά εικονοστοιχεία του πλαισίου είτε κατά γραμμές, είτε κατά στήλες. Από την παραπάνω διαδικασία προκύπτουν 6 γινόμενα τα οποία εν συνεχεία αθροίζονται και κανονικοποιούνται προκειμένου να παράξουν ένα υποεικονοστοιχείο (μεταξύ του 3 και 4 δείγματος του φίλτρου). Η διαδικασία επαναλαμβάνεται για κάθε “οριζόντιο” και “κατακόρυφο” ημι-εικονοστοιχείο μετακινώντας τον πυρήνα κατά μήκος των γραμμών και καθ’ ύψος των στηλών όλου του πλαισίου. Επιπρόσθετα, εφαρμόζοντας τον ανωτέρω πυρήνα στις ήδη υπολογισμένες τιμές των οριζόντιων και κατακόρυφων ημι-εικονοστοιχείων υπολογίζονται και οι τιμές των “διαγώνιων” ημι-εικονοστοιχείων. Συνεπώς, ανάλογα με την τελική θέση του δείγματος εικόνας απαιτείται η επεξεργασία από 6 έως 36 εικονοστοιχείων για την παραγωγή ενός ημι-εικονοστοιχείου.

Προκειμένου να αποφύγουμε το δαπανηρό φίλτρο του H.264 στη μονάδα ΕΚ, σε αυτό το κεφάλαιο μελετούμε παρόμοιες τεχνικές παρεμβολής οι οποίες χρησιμοποιούν λιγότερα από 6 δείγματα [18] [46] [14] και παράλληλα εκμεταλλευόμα-

στε τις ιδιότητες της κλίσης της φωτεινότητας επί της εικόνας [82].

3.1 Τεχνικές Παρεμβολής

Στην παρούσα ενότητα παρουσιάζονται αρχικά τρεις ήδη γνωστές τεχνικές και εν συνεχεία εισαγάγονται τρεις νέες τεχνικές οι οποίες έχουν ως στόχο να βελτιώσουν τη διαδικασία παρεμβολής.

Θεωρούμε ότι κάθε εικονοστοιχείο της αρχικής εικόνας βρίσκεται εντός ενός δισδιάστατου πλέγματος στη διάκριτη θέση (i, j) , όπου με $i, j \in \mathbb{N}$ συμβολίζουμε αντίστοιχα τις κατακόρυφες και οριζόντιες συντεταγμένες. Τα υπο-εικονοστοιχεία παράγονται στις θέσεις με συντεταγμένες $(i+k, j+l)$ όπου $k, l \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$. Να σημειώσουμε εδώ τη διάκριση μεταξύ των υπο-εικονοστοιχείων σε υποτετραπλάσια (quarter-) και υποδιπλάσια (half-) απόσταση, για τα οποία $k, l \notin \{\frac{1}{4}, \frac{3}{4}\}$. Τα υπο-εικονοστοιχεία σε υποδιπλάσια απόσταση, χωρίζονται με τη σειρά τους σε ημι-οριζόντια (ΗΟ ή Half-Horizontal - HH), ημι-κατακόρυφα (ΗΚ ή Half-Vertical - HV) και ημι-διαγώνια (ΗΔ ή Half-Diagonal - HD) τα οποία έχουν συντεταγμένες $(i+\frac{1}{2}, j+\frac{1}{2})$. Στην εικόνα 3.1 απεικονίζεται μία τυχαία περιοχή του αρχικού πλαισίου η οποία μεγεθύνεται προκειμένου να καταδείξει τις θέσεις όλων των υποεικονοστοιχείων τα οποία αφορούν το πρότυπο H.264. Επιπρόσθετα, στην ίδια εικόνα επισημαίνουμε θέσεις και περιοχές οι οποίες θα μας χρειαστούν στο υπόλοιπο του κεφαλαίου.

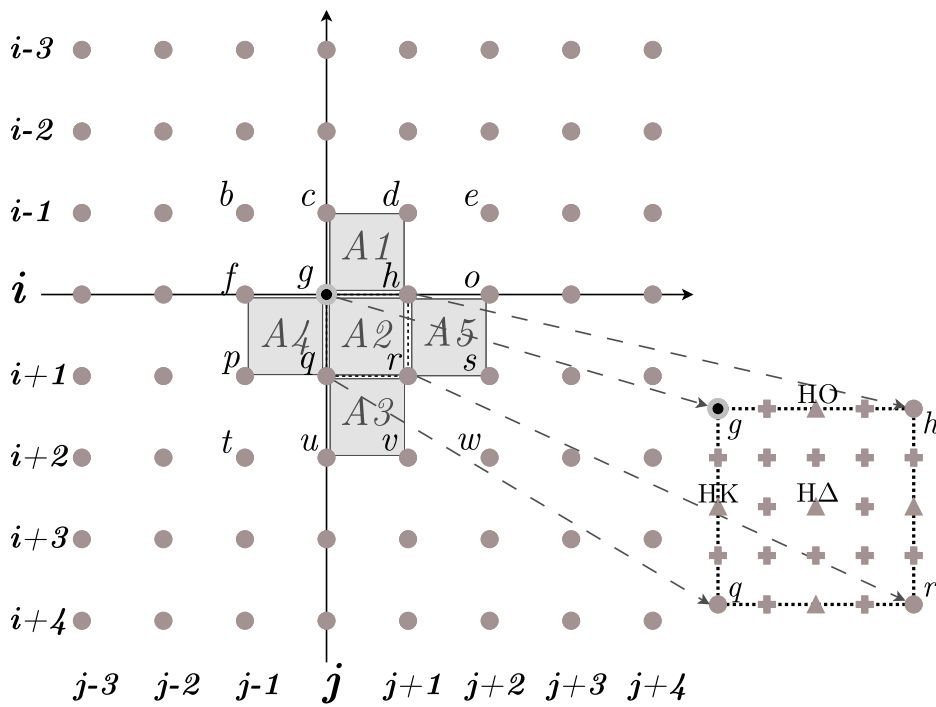
Κάθε ημι-εικονοστοιχείο παράγεται από μία διαδικασία παρεμβολής η οποία εφαρμόζεται στη γειτονική περιοχή της θέσης που μας ενδιαφέρει. Οι διαδικασίες τις οποίες μελετούμε, είναι οι εξής:

3.1.1 Διγραμμική Παρεμβολή

Στην πράξη πρόκειται για τον απλό μέσο όρο δύο παρακείμενων εικονοστοιχείων του αρχικού πλαισίου τα οποία περικλείουν το ΗΟ ή ΗΚ εικονοστοιχείο προς υπολογισμό [18]. Για τον υπολογισμό του ΗΔ εικονοστοιχείου, υπολογίζουμε το μέσο όρο των 4 γειτονικών εικονοστοιχείων $\{g, h, q, r\}$ (Εικ. 3.1).

3.1.2 Δικυβική Παρεμβολή

Η τεχνική αυτή βασίζεται στην επίλυση πολυωνύμων τρίτης τάξης [46]. Εμείς εξετάζουμε την παραμετροποιημένη μορφή της υποκείμενης εξίσωσης χρησιμοποιώντας την παράμετρο $a \in [-1, 0]$ προκειμένου να μεταβάλλουμε την “οξύτητα” της προκύπτουσας εικόνας. Κατά τη μελέτη μας επικεντρωθήκαμε στις τιμές $a = -1$, $a = -0.75$, και $a = -0.5$ από τις οποίες προκύπτουν τρεις διάκριτοι πυρήνες παρεμβολής με τους παρακάτω συντελεστές αντίστοιχα: $\langle -1, 5, 5, -1 \rangle$, $\langle -3, 19, 19, -3 \rangle$ και $\langle -1, 9, 9, -1 \rangle$. Κάθε τετράδα συντελεστών πολλαπλασιάζεται με 4 διαδοχικά εικονοστοιχεία για να παράξει το αντίστοιχο ενδιάμεσο ημεικο-



Εικόνα 3.1: Πλέγμα εικονοστοιχείων εικόνας και απεικόνιση θέσεων υποεικονοστοιχείων με μεγέθυνση περιοχής διαστάσεων 1×1 (δεξιά). Ο συμβολισμός επελέγη ώστε να διευκολύνει την περιγραφή των φίλτρων.

νοστοιχείο. Κατά τη δικυβική παρεμβολή, το ΗΔ εικονοστοιχείο προκύπτει από την εφαρμογή του πυρήνα επάνω σε 4 ήδη υπολογισμένα ημεικονοστοιχεία.

3.1.3 Lanczos Παρεμβολή

Όντας παρόμοια υπολογιστικά με την παρεμβολή του προτύπου H.264/AVC, η παρεμβολή Lanczos τρίτης τάξης χρησιμοποιεί ένα φίλτρο 6 δειγμάτων. Η τεχνική βασίζεται στη συνάρτηση ημιτόνου ανηγμένου στη γωνία του (Cardinal Sine, Sinc) [14]. Εδώ εξετάζουμε τον πυρήνα με τους εξής προϋπολογισμένους συντελεστές: $\langle \frac{12}{50\pi^2}, -\frac{12}{9\pi^2}, \frac{6}{\pi^2}, \frac{6}{\pi^2}, -\frac{12}{9\pi^2}, \frac{12}{50\pi^2} \rangle$.

3.1.4 Τριγωνισμός Εξαρτώμενος από τα Δεδομένα εισόδου (ΤΕΔ, Data-Dependent Triangulation - DDT)

Επί του παρόντος προτείνουμε μια τροποποίηση της τεχνικής η οποία προτάθηκε στο [82]. Οι συγγραφείς του [82] βασίζονται στη διαδικασία ανίχνευσης ακμών προκειμένου να καταλήξουν στο σύνολο των αρχικών εικονοστοιχείων τα οποία θα εισαχθούν στη συνάρτηση παρεμβολής. Εμείς μελετούμε μια ειδική περίπτωση του ΤΕΔ κατά την οποία εξετάζουμε μόνο 4 εικονοστοιχεία. Για να διευκολύνουμε την περιγραφή της τεχνικής, συμβολίζουμε το ΗΟ εικονοστοιχείο ως Y_D^H στη θέση $(i, j + \frac{1}{2})$ και το ΗΚ εικονοστοιχείο ως Y_D^{HV} στη θέση $(i + \frac{1}{2}, j)$ (Εικ. 3.1). Προκειμένου να αποφασίσουμε εάν τα εικονοστοιχεία $\{g, h, q, r\}$ περικλείουν κάποια ακμή, υπολογίζουμε τις διαφορές φωτεινότητάς τους ως εξής: εάν $|Y_g - Y_r| > |Y_h - Y_q|$, τότε θεωρούμε ότι έχουμε ανιχνεύσει ακμή στο hq , διαφορετικά η ακμή βρίσκεται στο rg . Στην πρώτη περίπτωση - την οποία συμβολίζουμε ως gE_D - υποθέτουμε ότι τα εικονοστοιχεία $\{g, h, q\}$ σχηματίζουν μια ομοιογενή τριγωνική περιοχή και υπολογίζουμε τις τιμές φωτεινότητας:

$$\begin{aligned} Y_D^{HH} &= \text{Clip}_{div_D}^R(w_1 Y_g + w_1 Y_h + w_2 Y_q) \\ Y_D^{HV} &= \text{Clip}_{div_D}^R(w_1 Y_g + w_1 Y_q + w_2 Y_h) \end{aligned} \quad (3.1)$$

όπου με $\text{Clip}_{div_D}^R$ συμβολίζουμε μία συνάρτηση κανονικοποίησης (η τελευταία διαιρεί με $div_D = 2w_1 + w_2$, και περιορίζει την τιμή στο διάστημα $[0, 255]$). Μέσω των συντελεστών $w_1 > w_2$ ρυθμίζουμε τα βάρη της φωτεινότητας των γειτόνων του υποεικονοστοιχείου προς υπολογισμό. Κατόπιν εκτενούς μελέτης, η υψηλότερη τιμή PSNR μετρήθηκε για συντελεστές $w_1 = 7$ and $w_2 = 2$ (δεδομένου ότι παράγουν παρονομαστή ο οποίος είναι δύναμη του 2: $div_D = 2^4$). Παρόμοια διαδικασία (προσανατολισμός και βάρη) ακολουθούμε και εφόσον η ακμή ανιχνευθεί στη διεύθυνση rg (δηλαδή όταν συμβολικά ισχύει gE_D). Σε αυτή την περίπτωση τροποποιούμε κατάλληλα τις τιμές φωτεινότητας της Εξ. 3.1. Η τεχνική εκφυλίζεται σε μία απλή διγραμμική παρεμβολή στην περίπτωση ομοιογενούς τετραγώνου στο $ghqr$ (οπότε και $w_1=1, w_2=0$).

Για τον υπολογισμό του ΗΔ εικονοστοιχείου, προτείνουμε έναν επιπρόσθετο έλεγχο της κλίσης, ο οποίος έπεται της ανίχνευσης gE_D ή rE_D ακμής. Ο υπολογισμός αυτός στοχεύει στην ανίχνευση του πιο ομοιογενούς τριγώνου στην

περιοχή A_2 (βλέπε Εικ. 3.1). Συνεπώς, στην περίπτωση ${}^q E_D$ ακμής ελέγχουμε την τιμή $|Y_g - Y_q| + |Y_g - Y_h| < |Y_r - Y_q| + |Y_r - Y_h|$, ειδάλτως την τιμή $|Y_h - Y_g| + |Y_h - Y_r| < |Y_q - Y_g| + |Y_q - Y_r|$ προκειμένου να αποφασίσουμε αν το ΗΔ εικονοστοιχείο θα βρίσκεται επάνω (*above* $<$) ή κάτω (*below* $>$) από την ακμή αυτή. Στο σημείο αυτό επεκτείνουμε τον ανωτέρω συμβολισμό με τους υπερδείκτες *abu* και *blw* για να περιγράψουμε τον τροποποιημένο ΤΕΔ (τΤΕΔ) ως:

$$Y_D^{HD} = \begin{cases} \text{Clip}_{div_D}^{\mathcal{R}}(w_1 Y_h + w_1 Y_q + w_2 Y_g) & \text{όταν } {}^q E_D^{abv} \\ \text{Clip}_{div_D}^{\mathcal{R}}(w_1 Y_h + w_1 Y_q + w_2 Y_r) & \text{όταν } {}^q E_D^{blw} \\ \text{Clip}_{div_D}^{\mathcal{R}}(w_1 Y_g + w_1 Y_r + w_2 Y_h) & \text{όταν } {}^g E_D^{abv} \\ \text{Clip}_{div_D}^{\mathcal{R}}(w_1 Y_g + w_1 Y_r + w_2 Y_q) & \text{όταν } {}^g E_D^{blw} \\ \text{Clip}_4^{\mathcal{R}}(Y_g + Y_h + Y_q + Y_r) & \text{ομοιογένεια} \end{cases} \quad (3.2)$$

όπου τα βάρη w_1, w_2 και η συνάρτηση $\text{Clip}_{div_D}^{\mathcal{R}}$ είναι όπως περιγράφονται στην Εξ. 3.1.

Διακρίνουμε επίσης μια εναλλακτική προσέγγιση κατά την οποία χρησιμοποιούμε την Εξ. 3.1 για να περιγραφεί μια απλούστερη διαδικασία υπολογισμού των ΗΔ εικονοστοιχείων την οποία συμβολίζουμε ως τΤΕΔ'. Η τελευταία εφαρμόζει Διγραμμική παρεμβολή επί των δύο εικονοστοιχείων της ακμής τα οποία προέκυψαν κατά τον πρώτο έλεγχο του ΤΕΔ: $Y_{D'}^{HD} = \text{Clip}_2^{\mathcal{R}}(Y_h + Y_q)$ όταν ${}^q E_D$.

3.1.5 Ακμοστρεφής τεχνική CrossHD

Η παρούσα προτεινόμενη ακμοστρεφής τεχνική βασίζεται στη διαφορά φωτεινότητας περιοχών αντί μεμονωμένων εικονοστοιχείων, έχοντας ως στόχο τη βελτίωση των αποτελεσμάτων ανίχνευσης του ΤΕΔ. Αρχικά προσθέτουμε τη φωτεινότητα τεσσάρων γειτονικών εικονοστοιχείων για να υπολογίσουμε τη φωτεινότητα της μικρής τετραγωνικής περιοχής την οποία αυτά ορίζουν. Για παράδειγμα, υπολογίζουμε την τιμή $Y_{A1} = Y_c + Y_d + Y_g + Y_h$ (Εικ. 3.1). Με την τεχνική μας αυτή ελέγχουμε τις διαφορές $|Y_{A4} - Y_{A5}| > |Y_{A1} - Y_{A3}|$ και αποφασίζουμε αν η περιοχή A_2 περιέχει μια κατακόρυφη ($>$) ή οριζόντια ($<$) ακμή. Στην πρώτη περίπτωση, εξετάζουμε μεμονωμένα τις περιοχές A_1 , A_2 , και A_3 και μέσω ενός ελέγχου ΤΕΔ συμπεραίνουμε τις κατευθύνσεις των ακμών οι οποίες διέρχονται από κάθε περιοχή. Η διεύθυνση της πλειοψηφίας των ακμών εντός των περιοχών A_1 , A_2 , και A_3 καθορίζει και τη διεύθυνση εντός της περιοχής A_2 δηλαδή συμπεραίνουμε εάν πρόκειται για ακμή ${}^q E_x$ ή ${}^g E_x$. Όταν η αρχική ακμή έχει οριζόντια διεύθυνση, ελέγχουμε αντίστοιχα τις περιοχές A_4 , A_2 , και A_5 . Τέλος, υπολογίζουμε το ΗΔ εικονοστοιχείο από το μέσο όρο της φωτεινότητας των εικονοστοιχείων της ανιχνευθείσας ακμής: $Y_x^{HD} = \text{Clip}_2^{\mathcal{R}}(Y_h + Y_q)$ όταν ${}^q E_x$, ή $Y_x^{HD} = \text{Clip}_2^{\mathcal{R}}(Y_r + Y_g)$ όταν ${}^g E_x$. Στην περίπτωση που δεν ανιχνεύσουμε κάποια ακμή (δηλαδή η περιοχή A_2 χαρακτηρίζεται από ομοιογένεια), υπολογίζουμε το μέσο όρο φωτεινότητας των $\{g, h, q, r\}$.

3.1.6 Ακμοστρεφής τεχνική CxScale

Βασιζόμενοι στις ιδέες των τεχνικών που περιγράψαμε στις προηγούμενες υποενότητες, αναπτύσσουμε μία νέα τεχνική με την οποία βελτιώνουμε τόσο την ανίχνευση της ακμής, όσο και την επιλογή του επακόλουθου πυρήνα συνέλιξης. Εν προκειμένω, κατά την ανίχνευση ακμών εξετάζουμε 8 γειτονικά εικονοστοιχεία της αρχικής εικόνας και υπολογίζουμε εν συνεχεία τα ημεικονοστοιχεία με συνθηκοταγή χρήση Διγραμμικής ή Δικυβικής παρεμβολής. Η παρούσα τεχνική αποτελείται από τρία βήματα: i) ανίχνευση ακμής σε οριζόντια ή κατακόρυφη διεύθυνση, ii) ενδεχόμενη βελτίωση της ανιχνευθείσας κατεύθυνσης προς κάποια διαγώνιο και iii) επιλογή των εικονοστοιχείων για την εφαρμογή Διγραμμικής ή Δικυβικής παρεμβολής. Οι λεπτομέρειες εφαρμογής των βημάτων εξαρτώνται από τη θέση του ημι-εικονοστοιχείου προς υπολογισμό. Για τα εικονοστοιχεία σε ΗΟ θέση, η επαλήθευση της ανίσωσης $|(Y_f+Y_g)-(Y_h+Y_o)| < |(Y_c+Y_d)-(Y_q+Y_r)|$ μας καθορίζει την ύπαρξη ακμής σε οριζόντια διεύθυνση (${}^h E_C^{HH}$). Όταν στην προηγούμενη ανίσωση ισχύει “>” τότε η ακμή έχει κατακόρυφη διεύθυνση και επαναπροσδιορίζουμε την τελική κατεύθυνσή της από τους εξής ελέγχους:

περίπτωση ${}^d E_C^{HH}$ όταν $|Y_c - Y_r| > |Y_d - Y_q|$ (από q προς d)
περίπτωση ${}^c E_C^{HH}$ όταν $|Y_c - Y_r| < |Y_d - Y_q|$ (από r προς c)
περίπτωση ${}_{A1}^3 E_C^{HH}$ όταν $|Y_c - Y_r| = |Y_d - Y_q|$ (αυστηρά κατακόρυφη διεύθυνση)

Σε διαφορετική περίπτωση (όταν $|(Y_f+Y_g)-(Y_h+Y_o)| = |(Y_c+Y_d)-(Y_q+Y_r)|$) συμπεραίνουμε ότι πρόκειται για ομοιογενή περιοχή. Για τον υπολογισμό της φωτεινότητας, βασιζόμαστε στη σχέση:

$$Y_C^{HH} = \begin{cases} \text{Clip}_{32}^{\mathcal{R}}(-3Y_f + 19Y_g + 19Y_h - 3Y_o) & \text{όταν } {}^h E_C^{HH} \\ \text{Clip}_{32}^{\mathcal{R}}(-3Y_c + 19Y_d + 19Y_q - 3Y_r) & \text{όταν } {}^d E_C^{HH} \\ \text{Clip}_{32}^{\mathcal{R}}(-3Y_d + 19Y_c + 19Y_r - 3Y_q) & \text{όταν } {}^c E_C^{HH} \\ \text{Clip}_2^{\mathcal{R}}(Y_g + Y_h) & \text{διαφορετικά} \end{cases} \quad (3.3)$$

Κατά παρόμοιο τρόπο, για τις ΗΚ θέσεις η ανίσωση $|(Y_c+Y_g)-(Y_q+Y_u)| < |(Y_f+Y_p)-(Y_h+Y_r)|$ μας οδηγεί σε συμπέρασμα κατακόρυφης ακμής, ήτοι ${}^g E_C^{HV}$. Όταν η ακμή έχει οριζόντια διεύθυνση (>), επαναπροσδιορίζουμε την τελική της κατεύθυνση και υπολογίζουμε την τιμή φωτεινότητας Y_C^{HV} βάσει των:

περίπτωση ${}^h E_C^{HH}$ όταν $|Y_f - Y_r| > |Y_p - Y_h|$ (από p προς h)
περίπτωση ${}^r E_C^{HH}$ όταν $|Y_f - Y_r| < |Y_p - Y_h|$ (από f προς r)
περίπτωση ${}_{A4}^5 E_C^{HH}$ όταν $|Y_f - Y_r| = |Y_p - Y_h|$ (αυστηρά οριζόντια κατεύθυνση)

$$Y_C^{HV} = \begin{cases} \text{Clip}_{32}^{\mathcal{R}}(-3Y_c + 19Y_g + 19Y_q - 3Y_u) & \text{όταν } {}^g E_C^{HV} \\ \text{Clip}_{32}^{\mathcal{R}}(-3Y_f + 19Y_h + 19Y_p - 3Y_r) & \text{όταν } {}^h E_C^{HV} \\ \text{Clip}_{32}^{\mathcal{R}}(-3Y_h + 19Y_f + 19Y_r - 3Y_p) & \text{όταν } {}^r E_C^{HV} \\ \text{Clip}_2^{\mathcal{R}}(Y_g + Y_q) & \text{διαφορετικά} \end{cases} \quad (3.4)$$

Τέλος, για τον υπολογισμό των εικονοστοιχείων στις ΗΔ θέσεις εξετάζουμε

την ανίσωση $|(Y_b + Y_g) - (Y_r + Y_w)| > |(Y_e + Y_h) - (Y_q + Y_t)|$ από όπου προκύπτει ακμή ${}^h E_c^{HD}$. Διαφορετικά, υποθέτουμε ότι η ακμή βρίσκεται μεταξύ των $\{g, r\}$, ήτοι ${}^g E_c^{HD}$. Έπειτα, υπολογίζουμε τη φωτεινότητα ως:

$$Y_c^{HD} = \begin{cases} \text{Clip}_{32}^{\mathcal{R}}(-3Y_e + 19Y_h + 19Y_q - 3Y_t) & \text{if } {}^h E_c^{HD} \\ \text{Clip}_{32}^{\mathcal{R}}(-3Y_b + 19Y_g + 19Y_r - 3Y_w) & \text{if } {}^g E_c^{HD} \end{cases} \quad (3.5)$$

3.2 Αξιολόγηση τεχνικών

Για την αξιολόγηση των επιδόσεων των ανωτέρω τεχνικών εκτελούμε διάφορα σενάρια ΕΚ ακολουθούμενα από την πρότυπη διαδικασία ΑΚ. Σε κάθε σενάριο ΕΚ χρησιμοποιούμε και από μία διαφορετική τεχνική προκειμένου να ανιχνεύσουμε κλασματικές μετατοπίσεις πλακιδίων. Η δε διαδικασία της αντιστάθμισης βασίζεται εξ ολοκλήρου στα προκύπτοντα ανύσματα κίνησης και συνθέτει πλαίσια πρόβλεψης βάσει του προτύπου φίλτρου 6 δειγμάτων. Συνεπώς, με το χρησιμοποιούμενο σχήμα κωδικοποίησης εξασφαλίζουμε ότι κωδικοποιητής και αποκωδικοποιητής χρησιμοποιούν σε κάθε περίπτωση ταυτόσημα πλαίσια αναφοράς προκειμένου να παράξουν την πρόβλεψή τους. Με τον τρόπο αυτό αποφεύγεται η συσσώρευση σφαλμάτων από πιθανές “ολισθήσεις” μεταξύ κωδικοποιητή και αποκωδικοποιητή. Πιο συγκεκριμένα, ο αλγόριθμος ΕΚ συγκρίνει τα υποψήφια πλακίδια 4×4 εικονοστοιχείων μέσω του Αθροίσματος Απολύτων Διαφορών (ΑΑΔ, Sum of Absolute Differences - SAD) και αποφασίζει σε δύο φάσεις: i) αρχικά μία αναζήτηση με μοτίβο ρόμβου (“Diamond Search”) ταιριάζει το πλακίδιο στη βέλτιστη θέση ακέραιας μετατόπισης, ii) έπειτα ακολουθεί μία εξαντλητική αναζήτηση 8 υποψήφιων πλακιδίων γειτονικά της ευρεθείσας θέσης του προηγούμενου βήματος με την οποία ανιχνεύονται κλασματικές μετατοπίσεις σε απόσταση $\pm \frac{1}{2}$ εικονοστοιχείων. Κατά την αξιολόγησή μας, σε κάθε εκτέλεση του ανωτέρω αλγορίθμου χρησιμοποιούμε και από μία διαφορετική τεχνική παρεμβολής. Η αποδοτικότητα της τελευταίας καθορίζει και την ποιότητα των τελικών εικονορροών. Παρακάτω παραθέτουμε τα αποτελέσματα της αξιολόγησης στα οποία για κάθε τεχνική εμφανίζεται το ΔPSNR των τελικών εικονορροών.

Οι εξομοιώσεις μας περιλαμβάνουν διάφορες γνωστές εικονορροές (*foreman*, *car-phone*, *container*, *pedestrian*, *riverbed*, *rush-hour* κ.λ.π.) σε 5 διαφορετικές αναλύσεις. Στον πίνακα 3.1 παρατίθενται τα αποτελέσματα της μετρικής PSNR για τον ανωτέρω αλγόριθμο, όταν εξετάζονται κλασματικές μετατοπίσεις μόνο κατά την οριζόντια και κατακόρυφη διεύθυνση. Παρουσιάζονται αποτελέσματα από δύο φίλτρα 6 δειγμάτων (H.264/AVC, Lanczos), τρία φίλτρα 4 δειγμάτων (Δικυβική παρεμβολή) και δύο ακμοστρεφείς τεχνικές (TEΔ, CxScale). Για λόγους πληρότητας, ο πίνακας 3.1 περιλαμβάνει και τα χαμηλής ποιότητας αποτελέσματα της τεχνικής Πλησιέστερου Γείτονα (Nearest Neighbor - NN [18]). Η τελευταία απλώς προωθεί την τιμή του εικονοστοιχείου από τη θέση ακέραιας μετατόπισης στη γειτονική ΗΟ/ΗΚ θέση (ουσιαστικά παραλείπεται η ανίχνευση κλασματικών μετατοπίσεων).

Από τα αποτελέσματα της τεχνικής NN καταδεικνύεται ότι ακόμη και με μόλις

Πίνακας 3.1: Μετρική PSNR για το πρότυπο φίλτρο του H.264/AVC και ΔPSNR για τις υπόλοιπες τεχνικές όταν η ΕΚ αφορά αποκλειστικά ΗΟ+ΗΚ θέσεις (με αντιστάθμιση προτύπου H.264).

Μέθοδος: ανάλυση	H.264	Nearest Neighbor	Διχυβική		Lanczos	CxScale	TEΔ	
			a=-1	a=-0.5 a=-0.75				
QCIF	35.0379	-2.2069	-0.0142	-0.0214	-0.0105	0.0009	-0.3359	-0.2265
CIF	34.2930	-1.4229	-0.0150	-0.0340	-0.0166	-0.0042	-0.3697	-0.1994
SD1	33.1775	-0.5483	-0.0170	-0.0192	-0.0118	-0.0030	-0.2071	-0.1249
720p	32.3743	-0.3316	-0.0130	-0.0151	-0.0096	-0.0029	-0.1021	-0.0866
1080p	33.0837	-0.2084	-0.0122	-0.0172	-0.0123	-0.0042	-0.0810	-0.0697
μ.τ.	33.4971	-0.8843	-0.0144	-0.0209	-0.0120	-0.0027	-0.2116	-0.1372

4 υποψήφιος HO/HK θέσεις, η ποιότητα πρόβλεψης του αλγορίθμου βελτιώνεται έως και 2 dB (σε εικονορροές χαμηλής ανάλυσης). Παρατηρούμε επίσης ότι η χρήση του φίλτρου Lanczos 6 δειγμάτων παράγει προβλέψεις σχεδόν αντίστοιχης ποιότητας με το πρότυπο φίλτρο του H.264. Να σημειωθεί ότι μετά από εκτενείς πειραματισμούς, καταλήξαμε στην προσέγγιση των συντελεστών του φίλτρου Lanczos από τους ακεραίους $\langle 3, -17, 78, 78, -17, 3 \rangle$ οι οποίοι επιτρέπουν πράξεις χαμηλής πολυπλοκότητας. Οι επιδόσεις των υπολοίπων τεχνικών κυμαίνονται μεταξύ των δύο ανωτέρω ακραίων περιπτώσεων των φίλτρων 6 και 0 δειγμάτων. Μεταξύ αυτών, τα φίλτρα της Δικυβικής παρεμβολής παράγουν την υψηλότερη δυνατή ποιότητα, όπως φαίνεται και από τις μετρήσεις του πίνακα 3.1. Να σημειώσουμε εδώ ότι παρουσιάζονται οι τρεις τιμές της παραμέτρου για τις οποίες μετρήθηκε η μεγαλύτερη τιμή της μετρικής PSNR (έχοντας προηγουμένως εξετάσει τις εξής τιμές $-a \in \{\frac{7}{8}, \frac{6}{8}, \frac{5}{8}, \frac{4}{8}, \frac{3}{8}, \frac{2}{8}\}$). Όπως φαίνεται από τον πίνακα, στην πλειοψηφία των περιπτώσεων ο πυρήνας συνέλιξης με συντελεστές $\langle -3, 19, 19, -3 \rangle$ (για $a = -0.75$) παράγει την υψηλότερη ποιότητα, εφόσον περιορίζει την αναμενόμενη υποβάθμιση του PSNR σε περίπου 0.01 dB (υπενθυμίζουμε ότι συγκρίνουμε με το πρότυπο φίλτρο). Το αποτέλεσμα αυτό έρχεται σε αντίθεση με την παρατήρηση ότι ο πυρήνας $\langle -1, 5, 5, -1 \rangle$ αποτελεί -δαισθητικά- καλύτερη προσέγγιση του προτύπου πυρήνα $\langle 1, -5, 20, 20, -5, 1 \rangle$. Οι ακμοστρεφείς τεχνικές μειώνουν την ποιότητα κατά 0.1 dB, συνεπώς η επιφάνεια σφάλματος την οποία παράγουν αποκλίνει από εκείνη των φίλτρων 6 δειγμάτων. Παρόλα αυτά, όταν παραλείψουμε την πρότυπη αντιστάθμιση το PSNR των ακμοστρεφών τεχνικών υπερिशύει (έως και κατά 0.1 dB) ειδικά σε εικονορροές υψηλής ευκρίνειας. Όπως φαίνεται και από τον πίνακα 3.1 οι επιδόσεις των τεχνικών TEΔ και CxScale βελτιώνονται όσο αυξάνεται η ανάλυση του πλαισίου.

Εν συνεχεία εξετάζουμε την απόδοση των τεχνικών όταν υπολογίζονται εικονοστοιχεία σε 4 υποψήφιος HD θέσεις (για τα οποία οι υπολογιστικές απαιτήσεις είναι υψηλότερες εν συγκρίσει με αυτές των HO/HK θέσεων). Στον πίνακα 3.2 παρατίθενται τα αποτελέσματα της PSNR μετρικής των τεχνικών του πίνακα 3.1 μαζί με τέσσερις ακόμη ακμοστρεφείς τεχνικές: CrossHD, τTEΔ, τTEΔ' και της δημοσιευμένης στο [82] με Διγραμμική παρεμβολή στο τελευταίο στάδιο. Τα ποιοτικά συμπεράσματα που προκύπτουν είναι παρόμοια με αυτά του πίνακα 3.1, επιβεβαιώνοντας ότι με τη Δικυβική παρεμβολή και ειδικά τον πυρήνα συνέλιξης $\langle -3, 19, 19, -3 \rangle$ επιτυγχάνονται καλύτερα αποτελέσματα εν συγκρίσει με τις ακμοστρεφείς τεχνικές. Ωστόσο, παρατηρούμε ότι το τμήμα που είναι υπεύθυνο για τον υπολογισμό των HD εικονοστοιχείων της CxScale τεχνικής χρησιμοποιεί έναν αποδοτικό έλεγχο κλίσης και, σε συνδυασμό με τον πυρήνα της Δικυβικής παρεμβολής καθιστούν τη CxScale ως την επικρατέστερη μεταξύ των εξεταζομένων ακμοστρεφών τεχνικών. Όσον αφορά τις ακμοστρεφείς τεχνικές με φίλτρα μικρότερου μεγέθους, η τεχνική CrossHD ξεχωρίζει ανάμεσα στις υπόλοιπες προτεινόμενες παραπλήσιες με τον TEΔ τεχνικές.

Τέλος, εξετάζουμε τις επιδόσεις των τεχνικών, συνυπολογίζοντας όλες τις θέσεις ημεικονοστοιχείων για τις 8 πιθανές μετατοπίσεις κατά τη δεύτερη φάση της ΕΚ. Ο πίνακας 3.3 παραθέτει τα αποτελέσματα της PSNR μετρικής καθώς

Πίνακας 3.2: Μετρική PSNR για το πρότυπο φίλτρο του H.264/AVC και ΔPSNR για τις υπόλοιπες τεχνικές όταν η ΕΚ αφορά αποκλειστικά ΗΔ θέσεις (με αντιστάθμιση προτύπου H.264).

Μέθοδος: ανάλυση	H.264 Πλησιέστερου Γείτωνα	Δικυβική		Lanczos	CxScale	τTEΔ	[82]	CrossHD	τTEΔ/ τTEΔ'		
		a=-1	a=-0.5							a=-0.75	
QCIF	34.7318	-1.8864	-0.0288	-0.0436	-0.0143	0.0004	-0.1010	-0.1728	-0.1095	-0.1299	-0.1595
CIF	33.9850	-1.1102	-0.0145	-0.0423	-0.0148	-0.0016	-0.0740	-0.1731	-0.1219	-0.1414	-0.1595
SD1	33.1292	-0.4790	-0.0247	-0.0241	-0.0117	-0.0032	-0.0396	-0.1217	-0.0860	-0.0972	-0.1070
720p	32.3766	-0.3178	-0.0176	-0.0188	-0.0092	-0.0031	-0.0294	-0.0816	-0.0636	-0.0696	-0.0760
1080p	33.0869	-0.1979	-0.0146	-0.0223	-0.0119	-0.0045	-0.0420	-0.0746	-0.0589	-0.0637	-0.0725
μ.τ.	33.3785	-0.7512	-0.0202	-0.0292	-0.0121	-0.0004	-0.0495	-0.1220	-0.0864	-0.0984	-0.1121

Πίνακας 3.3: Μετρήσεις Ποιότητας και Χρόνου υπολογισμού κατά την Εκτίμηση Κίνησης σε HO+HK+HΔ θέσεις.

Τεχνική παρεμβολής	PSNR (dB)			Χρόνος (μsec) ανά πλακίδιο
	QCIF	SD1	1080p	
<i>H.264/AVC</i>	35.4263	33.3687	33.1513	46.0
<i>Lanczos</i>	-0.0032	-0.0050	-0.0076	46.0
<i>Δικυβική, $a=-0.75$</i>	-0.0215	-0.0177	-0.0202	30.6
$\text{TE}\Delta\oplus\tau\text{TE}\Delta$	-0.3513	-0.1782	-0.1018	21.3
$\text{TE}\Delta\oplus\tau\text{TE}\Delta'$	-0.3341	-0.1642	-0.0980	16.0
<i>CxScale</i>	-0.3801	-0.1798	-0.0904	45.6
$\text{TE}\Delta\oplus\text{CrossHD}$	-0.3192	-0.1618	-0.0932	32.4
$\text{TE}\Delta\oplus\text{CxScale}(\text{HD})$	-0.3061	-0.1302	-0.0839	28.3
$\text{TE}\Delta\oplus[82](\text{HD})$	-0.2913	-0.1492	-0.0889	53.6

και το χρόνο υπολογισμού για πλακίδια 16×16 ημεικονοστοιχείων σε αυθαίρετες θέσεις. Να σημειωθεί ότι έχουμε υπολογίσει το μέσο όρο για HO, HK και HΔ θέσεις και χρησιμοποιούμε το χρόνο εκτέλεσης σε έναν επεξεργαστή γενικού σκοπού αρχιτεκτονικής Core 2 x86-64 στα 3GHz ως μέτρο της πολυπλοκότητας (run-time complexity) της κάθε τεχνικής. Επίσης, όπως φαίνεται από τον ίδιο πίνακα, πραγματοποιούμε συνδυασμούς τεχνικών για HO/HK και HΔ θέσεις βασιζόμενοι στις επικρατέστερες ακμοστρεφείς τεχνικές των πινάκων 3.1 και 3.2. Στον πίνακα 3.3, χρησιμοποιούμε το συμβολισμό " $A \oplus B$ " για να περιγράψουμε το συνδυασμό της τεχνικής A σε HO/HK θέσεις και της τεχνικής B σε HΔ θέσεις. Συμπερασματικά, η χρήση Δικυβικής παρεμβολής ελαττώνει το χρόνο της συνέλιξης 6 δειγμάτων κατά 33% διατηρώντας την τιμή του PSNR σε παρόμοια επίπεδα (απόκλιση 0.02 dB). Οι βασισμένες σε TEΔ τεχνικές ελαττώνουν το χρόνο κατά 65% (κυρίως λόγω του ταχέως υπολογισμού των HΔ θέσεων) κοστίζοντας περίπου 0.1 dB. Από τα αποτελέσματα φαίνεται επίσης το υψηλό υπολογιστικό κόστος των ελέγχων όσον αφορά στη CxScale και την τεχνική του [82]. Ωστόσο, ο συνδυασμός των HΔ ελέγχων της CxScale με τους HO/HK ελέγχους του TEΔ έχουν ως αποτέλεσμα μία υβριδική τεχνική η οποία χαρακτηρίζεται από το υψηλότερο PSNR μεταξύ των ακμοστρεφών τεχνικών και βελτιωμένο χρόνο υπολογισμού κατά 40%.

ΚΕΦΑΛΑΙΟ 4

Υπερ-ανάλυση από Μία Εικόνα με τη χρήση Χαμηλής Πολυπλοκότητας Προσαρμοστικής Επαναλαμβανόμενης Οπισθοπροβολής

Η τελευταία δεκαετία, χαρακτηρίζεται από διαρκή και αξιοσημείωτη εξέλιξη στους τομείς της σχεδίασης μοντέλων και αλγορίθμων. Παρόλα αυτά, η έρευνα στο πεδίο της ΥΑ εξακολουθεί να στοχεύει στην παροχή βελτιωμένων λύσεων σε δύο βασικά προβλήματα. Το πρώτο, είναι η μειωμένη στιβαρότητα των μεθόδων εξαιτίας της ανακριβούς μοντελοποίησης και το δεύτερο, το διαρκώς κλιμακούμενο υπολογιστικό κόστος. Στόχος μας σε αυτό το κεφάλαιο είναι η παροχή μιας αποτελεσματικής λύσης στο πρόβλημα του υπολογιστικού κόστους. Για το σκοπό αυτό, προτείνουμε την ελάττωση της πολυπλοκότητας σε μεθόδους ΥΑ με την απευθείας εφαρμογή μετασχηματισμών εικόνας όπως προέκυψαν από πειραματικά αποτελέσματα. Συγκεκριμένα, προτείνουμε μία ακμοπροσαρμοστική τεχνική οπισθοπροβολής σφάλματος, η οποία βελτιώνει την ποιότητα της ζητούμενης εικόνας υψηλής ευκρίνειας (ΥΕ, High-Resolution - HR).

4.1 Μοντέλο παρατήρησης εικόνων και ΕΟΠ από μία εικόνα

Στην ενότητα αυτή περιγράφουμε τη μοντελοποίηση της διαδικασίας απεικόνισης καθώς και τον αλγόριθμο ΕΟΠ από μία εικόνα όπως αυτός διατυπώθηκε στο [40]. Χρησιμοποιούμε κεφαλαία για το συμβολισμό των πινάκων και πεζά για το συμβολισμό των βαθμωτών.

Έστω ότι S_{HR} και O_{LR} είναι αντίστοιχα η ζητούμενη λύση ΥΕ (αρχική εικόνα) και η ΧΕ (Χαμηλής Ευκρίνειας, Low-Resolution - LR) παρατήρηση (δεδομένη εικόνα). Θεωρώντας ένα απλοποιημένο μοντέλο παρατήρησης εικόνων, η

παρατηρούμενη εικόνα ΧΕ προκύπτει από τον ΥΕ αντίστοιχό της ως:

$$\mathcal{O}_{LR} = \mathcal{D}(\mathcal{A}(\mathcal{S}_{HR})) + \mathcal{E}_{LR} \quad (4.1)$$

όπου με \mathcal{A} συμβολίζουμε τον τελεστή αντιαναδίπλωσης (anti-aliasing, AA), με \mathcal{D} τον τελεστή αποδεκάτισης και με \mathcal{E}_{LR} την εικόνα σφάλματος η οποία αντιστοιχεί στην παρατήρηση ΧΕ. Η εξίσωση 4.1 μπορεί να γραφεί και ως

$$\mathcal{O}_{LR} = \downarrow_f (H * \mathcal{S}_{HR}) + \mathcal{E}_{LR} \quad (4.2)$$

όπου f είναι ο παράγοντας αποδεκάτισης, H είναι η μήτρα της Συνάρτησης Διασποράς Σημείου (ΣΔΣ - Point Spread Function, PSF) η οποία αντιπροσωπεύει τη διαδικασία φιλτραρίσματος και $*$ είναι το σύμβολο της συνέλιξης. Η επίλυση της Εξ. 4.2 μας παρέχει τη μήτρα σφάλματος:

$$\mathcal{E}_{LR} = \mathcal{O}_{LR} - \downarrow_f (H * \mathcal{S}_{HR}) \quad (4.3)$$

Η επίλυση της Εξ. 4.1 προκειμένου να καταλήξουμε στη μήτρα \mathcal{S}_{HR} θεωρείται ένα κακώς-τεθειμένο (ill-posed) πρόβλημα, ακόμη και αν υποθεθεί ότι έχουμε ένα απλοποιημένο μοντέλο παρατήρησης με πολλές διαφορετικές εικόνες ΧΕ για κάθε \mathcal{S}_{HR} . Ο αλγόριθμος ΕΟΠ προσεγγίζει τη λύση \mathcal{S}_{HR} μέσα από μία επαναληπτική διαδικασία η οποία ελαχιστοποιεί το σφάλμα μεταξύ της παρατηρηθείσας εικόνας ΧΕ και της εκτιμηθείσας εικόνας ΧΕ (της υποδειγματοληπτικής εκδοχής της ανακατασκευασμένης εικόνας ΥΕ, ήτοι \mathcal{R}_{HR}). Το σφάλμα ανακατασκευής υπολογίζεται με την αντικατάσταση της μήτρας \mathcal{S}_{HR} από την \mathcal{R}_{HR} στην Εξ. 4.3. Σε κάθε επανάληψη i του αλγορίθμου ΕΟΠ, το υπολογισθέν σφάλμα \mathcal{E}_{LR}^i προβάλλεται στη μήτρα \mathcal{R}_{HR}^i προκειμένου να βελτιώσει την ποιότητα της ανακατασκευασμένης εικόνας ΥΕ. Ξεκινώντας από μία αρχική εκτίμηση ΥΕ \mathcal{R}_{HR}^0 , η ΕΟΠ επαναλαμβάνεται έως ότου ικανοποιηθεί κάποια συνθήκη τερματισμού (π.χ. ελαχιστοποίηση σφάλματος) ή όταν ξεπεραστεί το μέγιστο επιτρεπόμενο όριο επαναλήψεων. Θεωρώντας λοιπόν μόνο μία μήτρα παρατήρησης ΧΕ για κάθε λύση ΥΕ, καταλήγουμε στη μήτρα \mathcal{R}_{HR} μετά από n επαναλήψεις ως [17]:

$$\begin{aligned} \mathcal{R}_{HR}^n &= \mathcal{R}_{HR}^{n-1} + \mathcal{A}^{-1}(\mathcal{D}^{-1}(\mathcal{E}_{LR}^{n-1})) \\ &= \mathcal{R}_{HR}^{n-1} + (G * (\uparrow_f (\mathcal{E}_{LR}^{n-1}))) \end{aligned} \quad (4.4)$$

όπου f είναι ο παράγοντας υπερ-δειγματοληψίας και G η μήτρα οπισθοπροβολής. Όπως παρουσιάζεται στο [17], επιλέγοντας προσεκτικά τον πίνακα οπισθοπροβολής και το φίλτρο αποδεκάτισης, η μήτρα \mathcal{R}_{HR}^n θα συγκλίνει στη ζητούμενη εικόνα καθώς $n \rightarrow \infty$. Σε απλοποιημένα μοντέλα παρατήρησης, η ζητούμενη εικόνα θα ταυτίζεται με τη μήτρα \mathcal{S}_{HR} .

4.2 Προτεινόμενος αλγόριθμος ΥΑ

Ο κεντρικός σκοπός της μελέτης μας είναι η ελάττωση της πολυπλοκότητας σε μεθόδους ΥΑ υψηλής ποιότητας. Για το σκοπό αυτό βασιζόμαστε στο περιγραφέν

σχήμα ΕΟΠ και προτείνουμε έναν αλγόριθμο στον οποίο οι μήτρες αποδεκτικής H και οπισθοπροβολής G έχουν σταθερές τιμές και περιορισμένες διαστάσεις. Επιπρόσθετα, βασιζόμαστε σε διάκριτες τεχνικές παρεμβολής για τον υπολογισμό της αρχικής προσέγγισης ΥΕ και την επακόλουθη υπερ-δειγματοληψία της μήτρας σφάλματος. Κάθε τεχνική έχει προσαρμοστεί στις απαιτήσεις του αντίστοιχου βήματος του αλγορίθμου προκειμένου να εξισορροπήσει ικανοποιητικά μεταξύ του χρόνου υπολογισμού και της ποιότητας των αποτελεσμάτων.

Αρχικά υπερ-δειγματοληπτούμε την παρατηρούμενη εικόνα ΧΕ μέσω ενός φίλτρου χαμηλής πολυπλοκότητας m δειγμάτων για να καταλήξουμε στην αρχική ανακατασκευή ΥΕ, \mathcal{R}_{HR}^0 . Ακολούθως, πραγματοποιείται συνέλιξη της μήτρας \mathcal{R}_{HR}^0 με μία δισδιάστατη μήτρα η οποία αντιστοιχεί σε γκαουσιανή $\Sigma\Delta\Sigma$ και έχει το ρόλο φίλτρου αντανάδιπλωσης. Η προκύπτουσα εικόνα $H * \mathcal{R}_{HR}^0$ αποδεκατίζεται κατά f και παράγει τη μήτρα \mathcal{R}_{LR}^0 , μία ανακατασκευασμένη έκδοση της αρχικής παρατήρησης ΧΕ. Σε αυτό το σημείο, αφαιρούμε την ανακατασκευή ΧΕ από την παρατήρηση ΧΕ υπολογίζοντας έτσι τη μήτρα σφάλματος ΧΕ, \mathcal{E}_{LR}^0 . Η μήτρα \mathcal{E}_{LR}^0 υπερ-δειγματοληπτείται κατά f μέσω της ακμοπροσαρμοστικής τεχνικής \mathcal{L} , η οποία και θεωρείται ως βασική παράμετρος του προτεινόμενου αλγορίθμου. Κατόπιν, πραγματοποιείται συνέλιξη μεταξύ της μήτρας σφάλματος ΥΕ, \mathcal{E}_{HR}^0 με τη σταθερή μήτρα οπισθοπροβολής G , η οποία επίσης θεωρείται ως βασική αλγοριθμική παράμετρος. Η μήτρα σφάλματος ΧΕ χρησιμοποιείται ως ενδείκτης ποιότητας της ανακατασκευασμένης εικόνας μέσω του υπολογισμού της l^1 norm. Ανάλογα με την τιμή της l^1 norm, η φιλτραρισμένη μήτρα σφάλματος προβάλλεται στην αρχική ανακατασκευή ΥΕ προκειμένου να παράξει τη μήτρα \mathcal{R}_{HR}^1 , ολοκληρώνοντας έτσι την πρώτη επανάληψη του αλγορίθμου. Ο κύκλος αποδεκτικής-υπερ-δειγματοληψίας επαναλαμβάνεται για έναν πεπερασμένο αριθμό επαναλήψεων, βελτιώνοντας δυνητικά σε κάθε επανάληψη την τρέχουσα ανακατασκευασμένη μήτρα \mathcal{R}_{HR}^i μέσω της τρέχουσας μήτρας σφάλματος \mathcal{E}_{HR}^i (την οποία έχουμε επεξεργαστεί εκ νέου με οπισθοπροβολή). Τέλος, ο αλγόριθμος περιλαμβάνει και ένα μηχανισμό πρόωρου τερματισμού μέσω του οποίου ελέγχουμε τον ενδείκτη ποιότητας σε κάθε επανάληψη. Η ακριβής περιγραφή του αλγορίθμου παρατίθεται στην Εικ. 4.1.

Να σημειωθεί ότι οι παράμετροι $\{G, H, \mathcal{K}, \mathcal{L}\}$ οι οποίες παρατίθενται στην Εικ. 4.1 διαδραματίζουν εξέχοντα ρόλο τόσο στο χρόνο εκτέλεσης, όσο και την ποιότητα της προκύπτουσας εικόνας ΥΕ. Σημαντικό τμήμα της εργασίας σε αυτό το κεφάλαιο έχει αφιερωθεί στη σχεδίαση, τη δοκιμή και τη λεπτομερή ρύθμιση των ανωτέρω παραμέτρων με σκοπό να βελτιώσουμε την απόδοση του μελετούμενου σχήματος ΥΑ. Οι ακόλουθες ενότητες περιγράφουν το σύνολο των τεχνικών παρεμβολής οι οποίες είτε αναπτύχθηκαν είτε εξετάστηκαν στη θέση των $\{\mathcal{K}, \mathcal{L}\}$, καθώς και την προσέγγισή μας για τον καθορισμό των παραμέτρων $\{G, H\}$.

4.2.1 Εξερεύνηση τεχνικών παρεμβολής

Ένα από τα κύρια χαρακτηριστικά του προτεινόμενου αλγορίθμου αποτελεί η δυνατότητα χρήσης μιας πληθώρας τεχνικών παρεμβολής για την υπερ-δειγματοληψία

Είσοδος: \mathcal{O}_{LR} // εικόνα χαμηλής ευκρίνειας
Έξοδος: \mathcal{R}_{HR} // εικόνα υψηλής ευκρίνειας
Παράμετροι: σταθερές $G, H \in \mathbb{R}^{5 \times 5}$ και $n, f \in \mathbb{N}$
Παράμετροι: αλγόριθμοι υπερ-δειγματοληψίας \mathcal{K}, \mathcal{L}

- 1: αρχικοποίηση: $maxerror \leftarrow \infty$
- 2: υπερ-δειγματοληψία \mathcal{O}_{LR} κατά f μέσω του συμβατικού αλγορίθμου \mathcal{K} και αποτέλεσμα \mathcal{R}_{HR}^0 :
 $\mathcal{R}_{HR}^0 \leftarrow \uparrow_f^{\mathcal{K}}(\mathcal{O}_{LR})$
- 3: **για** $i = 0$ **έως** $n - 1$ **εκτέλεσε**
- 4: συνέλιξη \mathcal{R}_{HR}^i με H και αποτέλεσμα $H * \mathcal{R}_{HR}^i$
- 5: αποδεκίαση ($H * \mathcal{R}_{HR}^i$) και αποτέλεσμα \mathcal{R}_{LR}^i :
 $\mathcal{R}_{LR}^i \leftarrow \downarrow_f(H * \mathcal{R}_{HR}^i)$
- 6: αφαίρεση \mathcal{R}_{LR}^i από \mathcal{O}_{LR} και αποτέλεσμα \mathcal{E}_{LR}^i :
 $\mathcal{E}_{LR}^i \leftarrow \mathcal{O}_{LR} - \mathcal{R}_{LR}^i$
- 7: υπερ-δειγματοληψία \mathcal{E}_{LR}^i κατά f μέσω του ακμοπροσαρμοστικού αλγορίθμου \mathcal{L} και αποτέλεσμα \mathcal{E}_{HR}^i :
 $\mathcal{E}_{HR}^i \leftarrow \uparrow_f^{\mathcal{L}}(\mathcal{E}_{LR}^i)$
- 8: συνέλιξη \mathcal{E}_{HR}^i με G και αποτέλεσμα $G * \mathcal{E}_{HR}^i$:
- 9: υπολογισμός σφάλματος e^i μέσω l^1 norm :
 $e^i \leftarrow \|\mathcal{E}_{LR}^i\|_1$
- 10: **εάν** $e^i < maxerror$ **τότε**
- 11: βελτίωση \mathcal{R}_{HR}^i μέσω $G * \mathcal{E}_{HR}^i$ και αποτέλεσμα \mathcal{R}_{HR}^{i+1} :
 $\mathcal{R}_{HR}^{i+1} \leftarrow \mathcal{R}_{HR}^i + G * \mathcal{E}_{HR}^i$
- 12: **ειδώλλως**
- 13: **επέστρεψε** \mathcal{R}_{HR}^i // πρόωρος τερματισμός
- 14: **ολοκλήρωση ελέγχου**
- 15: ανανέωση: $maxerror \leftarrow e^i$
- 16: **ολοκλήρωση βρόχου**
- 17: **επέστρεψε** \mathcal{R}_{HR}^n

Εικόνα 4.1: Ο προτεινόμενος αλγόριθμος ΥΑ.

της αρχικής εικόνας και της επακόλουθης μήτρας σφάλματος (Εικ. 4.1). Η χρήση διάκριτων τεχνικών για τις δύο αυτές διεργασίες μας επιτρέπει να βελτιστοποιήσουμε την ποιότητα και την πολυπλοκότητα του χρόνου εκτέλεσης, όπως θα παρουσιάσουμε στην ενότητα 4.3. Στην παρούσα μελέτη δίνουμε έμφαση σε ακμοστρεφείς τεχνικές χαμηλής πολυπλοκότητας, λόγω της ιδιαίτερης βαρύτητας η οποία διέπει τη διατήρηση της πληροφορίας των ακμών κατά την οπισθοπροβολή του σφάλματος [17]. Ανάμεσα στο σύνολο των εξεταζομένων τεχνικών ανήκουν γνωστές μέθοδοι παρεμβολής αλλά και οι σχεδιασμένες από εμάς ακμοπροσαρμοστικές *CxScale* και *CrossHD* [29]. Παρακάτω εξετάζουμε την απόδοση του αλγορίθμου στην πράξη, όταν χρησιμοποιείται κάθε μία από τις τεχνικές αυτές στη θέση των \mathcal{K} ή \mathcal{L} . Με τον τρόπο αυτό επιλέγουμε το βέλτιστο συνδυασμό παραμέτρων για την εκτέλεση του αλγορίθμου συνολικά.

Δικυβική παρεμβολή

Πρόκειται για το γνωστό φίλτρο πεπερασμένης χρονιστικής απόκρισης (ΠΚΑ, Finite Impulse Response - FIR) 4 δειγμάτων το οποίο απαιτεί έως και 16 εικονοστοιχεία όταν υπερ-δειγματοληπούμε κατά παράγοντα 4. Κατά τη μελέτη μας εξετάζουμε το πολυώνυμο για $a = -1$ [29] από το οποίο προκύπτουν οι συντελεστές συνέλιξης $\langle -1, 5, 5, -1 \rangle$. Οι συντελεστές αυτοί παρείχαν και τη μέγιστη αντικειμενική ποιότητα για τις χρησιμοποιηθείσες εικόνες αναφοράς.

Παρεμβολή Lanczos

Οι συντελεστές της υπερ-δειγματοληψίας Lanczos προκύπτουν από την παραθύρωση της συνάρτησης Sinc (Cardinal Sine). Κατά τη μελέτη μας πειραματιστήκαμε με τη συνάρτηση Lanczos τρίτης τάξης, από την οποία προκύπτουν φίλτρα ΠΚΑ 6 δειγμάτων με συντελεστές $\langle \frac{12}{50\pi^2}, -\frac{12}{9\pi^2}, \frac{6}{\pi^2}, \frac{6}{\pi^2}, -\frac{12}{9\pi^2}, \frac{12}{50\pi^2} \rangle$.

Φίλτρο 6 δειγμάτων H.264

Όπως περιγράψαμε και στα προηγούμενα κεφάλαια, το πρότυπο συμπίεσης H.264 [3] ορίζει κατά τη διαπλασιασική πρόβλεψη (inter-prediction) τη διαδικασία αντιστάθμιση κίνησης (AK) σε επίπεδο υποεικονοστοιχείων με την υπερ-δειγματοληψία πλακιδίων εικόνας κατά παράγοντα 4 σε οριζόντια και κατακόρυφη διεύθυνση. Αρχικά ο αριθμός των εικονοστοιχείων διπλασιάζεται και στις δύο διευθύνσεις μέσω ενός διαχωρίσιμου φίλτρου Wiener 6 δειγμάτων με συντελεστές $\langle 1, -5, 20, 20, -5, 1 \rangle$. Εν συνεχεία, το πλέγμα εικονοστοιχείων που προκύπτει υπερ-δειγματοληπτείται κατά παράγοντα 2 και στις δύο διευθύνσεις μέσω ενός απλού διγραμμικού φίλτρου. Κατά τη μελέτη στο παρόν κεφάλαιο, εστιάζουμε στο φίλτρο 6 δειγμάτων.

Φίλτρο 8 δειγμάτων H.265

Το πρότυπο συμπίεσης εικονοροών υψηλής αποδοτικότητας (High Efficiency Video Coding Standard - HEVC ή H.265) αποτελεί το διάδοχο του προτύπου H.264. Το μοντέλο δοκιμών που οδήγησε στην προτυποποίηση του H.265 χρησιμοποιεί τεχνικές παρεμβολής οι οποίες προήλθαν από την πλατφόρμα λογισμικού Key Technical Area (KTA, [2]). Η τελευταία αναπτύχθηκε ώστε να συγκεντρώσει και να θεμελιώσει όλες τις βελτιώσεις επάνω στο προηγούμενο πρότυπο H.264. Αρχικά, τα προσχέδια του προτύπου HEVC προέβλεπαν δύο βασικές διαμορφώσεις, τη διαμόρφωση υψηλής αποδοτικότητας (High Efficiency - HE) και τη διαμόρφωση χαμηλής πολυπλοκότητας (Low-Complexity - LC). Η διαμόρφωση υψηλής αποδοτικότητας προέβλεπε τη χρήση ενός φίλτρου 12 δειγμάτων με συντελεστές $\langle -1, 8, -16, 24, -48, 161, 161, -48, 24, -18, 8, -1 \rangle$ για τον υπολογισμό των τιμών φωτεινότητας στην αρχική υπερ-δειγματοληψία κατά παράγοντα 2. Από την άλλη πλευρά, η διαμόρφωση χαμηλής πολυπλοκότητας περιγράφει κατευθυντικές τεχνικές παρεμβολής βάσει των συντεταγμένων του εικονοστοιχείου που πρέπει να υπολογιστεί. Τα ημι-οριζόντια (HO) και ημικατακόρυφα (HK) εικονοστοιχεία υπολογίζονται με βάση το φίλτρο 6 δειγμάτων $\langle 3, -17, 78, 78, -17, 3 \rangle$ το οποίο, όπως είχαμε παρατηρήσει στο Κεφ. 3 θα μπορούσε να θεωρηθεί ως μία προσέγγιση χαμηλής πολυπλοκότητας του πυρήνα συνέλιξης Lanczos τρίτης τάξης [29]. Τα εικονοστοιχεία στις ΗΔ θέσεις υπολογίζονται μέσω ενός μη-διαχωρίσιμου φίλτρου 12 δειγμάτων. Για τον υπολογισμό των ημικατακόρυφων, η παρούσα μελέτη εξετάζει το φίλτρο 8 δειγμάτων $\langle -3, 12, -39, 158, 158, -39, 12, -3 \rangle$ το οποίο περιγράφεται στο [41].

Ακμοστρεφής τεχνική CxScale

Η ακμοπροσαρμοστική τεχνική αυτή προτάθηκε στο Κεφάλαιο 3. Υπενθυμίζουμε ότι για κάθε αρχικό εικονοστοιχείο εξετάζονται 8 γειτονικά εικονοστοιχεία με χρήση απλών ελέγχων κλίσης. Κατόπιν, η τεχνική επιλέγει μεταξύ πυρήνων συνέλιξης 2 ή 4 δειγμάτων ανάλογα με το χαρακτηρισμό της ανιχνευθείσας ακμής. Τα εικονοστοιχεία που βρίσκονται εντός της περιοχής μιας ακμής υπολογίζονται με τη βοήθεια φίλτρου Δικυβικής παρεμβολής στην οποία η παράμετρος a τίθεται ίση με -0.75 . Στον αντίποδα, τα εικονοστοιχεία που βρίσκονται εντός ομοιογενών περιοχών, υπολογίζονται με τη βοήθεια Διγραμμικής παρεμβολής. Εν προκειμένω, υιοθετούμε την τεχνική CxScale για τον υπολογισμό των εικονοστοιχείων σε HO και HK θέσεις κατά την υπερ-δειγματοληψία της μήτρας σφάλματος.

Ακμοστρεφής τεχνική CrossHD

Η τεχνική αυτή προτάθηκε επίσης στο Κεφάλαιο 3 ως βελτίωση της τεχνικής ΤΕΔ για την ανίχνευση ακμών. Συνίσταται από ελέγχους κλίσης σε θέσεις παρακείμενες του ζητούμενου εικονοστοιχείου. Συγκεκριμένα, εφόσον ανιχνευθεί η κατεύθυνση κάποιας τοπικής ακμής, η τελική φωτεινότητα του εικονοστοιχείου

υπολογίζεται μέσω Διγραμμικής παρεμβολής. Στην έρευνά μας προτείνουμε τη χρήση της CrossHD για τον υπολογισμό των ΗΔ εικονοστοιχείων της μήτρας σφάλματος ΥΕ, \mathcal{E}_{HR}^i .

4.2.2 Φίλτρο αποδεκτίσης και πυρήνας οπισθοπροβολής

Όπως αναφέραμε και στην Ενότητα 4.2, η ποιότητα των τελικών αποτελεσμάτων του αλγορίθμου εξαρτάται σημαντικά και από την επιλογή δύο ακόμη παραμέτρων: του φίλτρου αποδεκτίσης H και του πυρήνα οπισθοπροβολής G . Οι παράμετροι αυτές επιδρούν στην ακρίβεια του μετρούμενου σφάλματος μεταξύ της αρχικής παρατήρησης και της ανακατασκευασμένης εικόνας και τελικώς στο πόσο επωφελής κρίνεται κάθε επανάληψη των βημάτων 3-16 (Εικ. 4.1). Σκοπός της βελτίωσης (δηλαδή, του επαναληπτικού τμήματος του αλγορίθμου) είναι η “αντιστροφή” των αποτελεσμάτων του φίλτρου αποδεκτίσης, με άλλα λόγια η αντιστάθμιση της θόλωσης που εφαρμόστηκε κατά το σχηματισμό της εικόνας ΧΕ. Συνεπώς, η κατάλληλη επιλογή των παραμέτρων βελτιώνει την ακρίβεια του υπερ-δειγματοληπτημένου σφάλματος και, σταδιακά, την ποιότητα της εικόνας ΥΕ. Στην πράξη, οι παρακάτω τιμές των H and G προέκυψαν από διαδοχικές δοκιμές και συνεχείς βελτιώσεις. Κατά εξερεύνηση του χώρου λύσεων η μελέτη μας λαμβάνει υπόψιν και την πολυπλοκότητα του αλγορίθμου. Για το σκοπό αυτό αποφεύγουμε να αυξήσουμε την πολυπλοκότητα της συνέλιξης περιορίζοντας τις διαστάσεις των H και G σε 5×5 στοιχεία (περισσότερες λεπτομέρειες γύρω από αυτή την επιλογή αναφέρονται στο Κεφάλαιο 5).

Για την επιλογή του φίλτρου H , ακολουθούμε τη συμβατική προσέγγιση της συνέλιξης με γκαουσιανή $\Sigma\Delta\Sigma$ στο διάστημα $[-2, 2]$ και τυπική απόκλιση $\sigma = 0.6$. Ο υπολογισμός των τελικών τιμών του δισδιάστατου $H \in \mathbb{R}^{5 \times 5}$ υπολογίζεται στις θέσεις $\langle x, y \rangle$ μέσω της Εξίσωσης 4.5:

$$H_{x,y,\sigma} = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.5)$$

Για τη διαμόρφωση του πυρήνα συνέλιξης G , χρησιμοποιούμε τους ήδη υπολογισμένους συντελεστές του H και την ανίσωση 7 από το [40] προκειμένου να παράξουμε μία αρχική έκδοση του G στο πεδίο της συχνότητας. Εν συνεχεία μετασχηματίζουμε τον G στο χωρικό τομέα (spatial domain) πραγματοποιώντας παράλληλα εκτεταμένες δοκιμές χρήσης του αλγορίθμου. Υπενθυμίζουμε ότι η διατήρηση της πληροφορίας των ακμών είναι εξέχουσας σημασίας καθ’ όλη τη διαδικασία της οπισθοπροβολής, συνεπώς ρυθμίζουμε τον G ποιοτικά ώστε να προσεγγίζει γνωστούς ακμοενισχυτικούς πυρήνες (για παράδειγμα λαπλασιανή γκαουσιανής συνάρτησης - Laplacian of Gaussian). Τέλος, λαμβάνοντας και πάλι υπόψιν μας την πολυπλοκότητα της συνέλιξης, επιλέγουμε τους τελικούς συντελεστές του G ως ακεραίους και μάλιστα δυνάμεις του 2. Το αποτέλεσμα της ανωτέρω εξερεύνησης είναι ένας σταθερός πυρήνας συνέλιξης 5×5 δειγμάτων ο οποίος παρατίθεται στην Εξ. 4.6:

$$G = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & 2 & 2 & 2 & -1 \\ -1 & 2 & 8 & 2 & -1 \\ -1 & 2 & 2 & 2 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix} \quad (4.6)$$

4.3 Αξιολόγηση ποιότητας και πολυπλοκότητας

Προκειμένου να αξιολογήσουμε την απόδοση του προτεινόμενου αλγορίθμου, αναπτύξαμε μία πλατφόρμα υπερ-δειγματοληψίας σε γλώσσα C η οποία μας επιτρέπει να εξετάζουμε συνδυασμούς αλγορίθμων, φίλτρων και συναρτήσεων υπολογισμού σφάλματος. Η είσοδος του αλγορίθμου προέρχεται από εικόνες αναφοράς με ποικίλλο περιεχόμενο και αναλύσεις από QCIF έως 1080p (π.χ. *carphone-QCIF*, *coastguard-CIF*, *pedestrian-SD1*, *riverbed-720p*, *rush-1080p*, κ.λ.π.). Αρχικά, οι εικόνες αυτές υπόκεινται σε συνέλιξη με το H και αποδεκατίζονται κατά παράγοντα f ($f/2$ σε κάθε διάσταση), γεννώντας κατ' αυτόν τον τρόπο τις παρατηρηθείσες εικόνες ΧΕ οι οποίες αποτελούν και την είσοδο σε κάθε αλγόριθμο ΥΑ. Κατά την εξερεύνησή μας μεταβάλλουμε τον αλγόριθμο ΥΑ (ή τις παραμέτρους του) εντός της αναπτυχθείσας πλατφόρμας και παράγουμε εικόνες ΥΕ προς αξιολόγηση. Έπειτα, συγκρίνουμε το επίπεδο φωτεινότητας από κάθε παραχθείσα εικόνα ΥΕ με την αρχική εικόνα αναφοράς μέσω δύο αντικειμενικών μετρικών ποιότητας: α) της Κορυφής Λόγου Σήματος-Θορύβου (PSNR) και β) του Δείκτη Μέσης Δομικής Ομοιότητας (Mean Structural Similarity Index, MSSIM [93]). Η τελευταία συνδυάζει στοιχεία μετρικών αντικειμενικής ανάλυσης με χαρακτηριστικά του ανθρώπινου συστήματος όρασης. Σημειώνουμε ότι για τον υπολογισμό του δείκτη χρησιμοποιούμε παράθυρα 5×5 εικονοστοιχείων, τη ΣΔΣ H και τις παραμέτρους $K_1=0.01$ και $K_2=0.03$ από την Εξ. (13) του [93].

Η κεντρική ιδέα πίσω από την αξιολόγηση του αλγορίθμου έγκειται στο να μετρήσουμε την απόδοσή του για διάφορους συνδυασμούς των παρεμβολών \mathcal{K} και \mathcal{L} (Εικ. 4.1). Για το σκοπό αυτό χρησιμοποιούμε τις τεχνικές που περιγράψαμε στην Ενότητα 4.2.1: τη Δικυβική με $a=-1$, τη Lanczos τρίτης τάξης (στην οποία αναφερόμαστε ως Lcz), την παρεμβολή του H.264 (ως AVC), την παρεμβολή του H.265 (ως $HEVC$) και το συνδυασμό των τεχνικών CxScale και CrossHD (ως $CxHD$, με τη CxScale να χρησιμοποιείται σε HO/HK θέσεις και την CrossHD για τις ΗΔ). Ως μέτρο σύγκρισης της ποιότητας, συμπεριλαμβάνουμε επίσης α) την υλοποίηση της παρεμβολής Lanczos από το πρόγραμμα Matlab R2011b ($MLcz$), β) τη Δικυβική παρεμβολή και τέλος, γ) τον κλασικό αλγόριθμο ΕΟΠ όπως περιγράφεται στο [40] και έχει υλοποιηθεί στο [9] με τη μέθοδο $MLcz$ να παράγει την αρχική ανακατασκευασμένη εικόνα. Οι πίνακες 4.1 και 4.2 παρουσιάζουν τα πιο χαρακτηριστικά αποτελέσματα των ανωτέρω δοκιμών για $f = 4$ και $n = 10$. Κάθε στήλη αντιστοιχεί σε ένα συγκεκριμένο συνδυασμό $\{\mathcal{K}, \mathcal{L}\}$ για τον προτεινόμενο αλγόριθμο: η μέθοδος \mathcal{K} αναγράφεται στην αρχή της στήλης και

η μέθοδος \mathcal{L} κάτω από την \mathcal{K} . Τα αποτελέσματα των υπόλοιπων 3 αλγορίθμων αναγράφονται στις στήλες 2, 12 και 13. Στους πίνακες αυτούς καταγράφεται η τιμή των μετρικών PSNR και MSSIM για τη Δικυβική παρεμβολή (στήλη 2), ενώ για τους υπόλοιπους αλγορίθμους εμφανίζονται οι διαφορές Δ PSNR και Δ MSSIM σε σχέση με τις τιμές της στήλης 2.

Σύμφωνα με τα αποτελέσματα του πίνακα 4.1, το σχήμα οπισθοπροβολής βελτιώνει την τελική ποιότητα ακόμη και στην περίπτωση που έχουμε θέσει $K, L \leftarrow$ Δικυβική (με μόνη εξαίρεση τις εικόνες υψηλής ανάλυσης). Επιπρόσθετα, η χρήση της CxHD στη θέση του L βελτιώνει ακόμη περισσότερο την τιμή του PSNR σε όλες τις αναλύσεις. Βλέπουμε επίσης, ότι η παρεμβολή Lanczos τρίτης τάξης, παρόλο που βασίζεται σε φίλτρα μεγαλύτερου μεγέθους από τη Δικυβική παρεμβολή, υποβαθμίζει μερικώς την τελική ποιότητα. Η χρήση της CxHD στη θέση του L βελτιώνει και σε αυτή την περίπτωση τα αποτελέσματα (στήλες 5,6). Το φίλτρο 6 δειγμάτων του H.264 υπερτερεί των προηγούμενων καταλήγοντας σε βελτίωση του PSNR κατά περισσότερο από 0.5 dB. Η υψηλότερη τιμή PSNR παρατηρείται όταν θέσουμε $K \leftarrow$ HEVC και $L \leftarrow$ CxHD. Ο τελευταίος συνδυασμός επιτυγχάνει βελτίωση άνω του 1.6 dB για την εικόνα *blue-sky* (μεγέθους 1080p) και πάνω από 0.6 dB κατά μέσο όρο. Τα αποτελέσματα αντιστρέφονται όταν θέσουμε $K \leftarrow$ CxHD, επομένως, στο παρόν σχήμα η χρήση της τεχνικής CxHD είναι πιο σημαντική κατά την υπερ-δειγματοληψία της μήτρας σφάλματος.

Εν συγκρίσει με την MLcz παρεμβολή, η προταθείσα από τους Irani και Peleg μέθοδος ΕΟΠ παρέχει σημαντική βελτίωση ποιότητας όπως φαίνεται από την τελευταία στήλη των πινάκων. Για παράδειγμα, στην περίπτωση της εικόνας αναφοράς *rush hour* σε ανάλυση SD1, η υλοποίηση της μεθόδου ([9]) βελτιώνει το αποτέλεσμα κατά περισσότερο από 5 dB. Παρόλα αυτά, η αντικειμενική ποιότητα του προτεινόμενου σχήματος υπερσιχθεί και των δύο προαναφερθεισών μεθόδων - ακόμη και όταν προωθήσουμε το αποτέλεσμα της δικής μας Δικυβικής παρεμβολής ως είσοδο στο [9]. Συνολικά, οι παραπάνω παρατηρήσεις επιβεβαιώνονται και από την MSSIM μετρική, όπως φαίνεται από τον πίνακα 4.2.

Στην εικόνα 5.5 βλέπουμε την οπτική αναπαράσταση των αποτελεσμάτων 4 διαφορετικών μεθόδων για την εικόνα αναφοράς *carphone* σε QCIF ανάλυση. Το αποτέλεσμα της Δικυβικής παρεμβολής (Εικ. 4.2β') παρουσιάζει εμφανή αναδίπλωση εν συγκρίσει με τις υπόλοιπες μεθόδους. Η δε ανακατασκευή μέσω του κλασικού σχήματος ΕΟΠ ([40], Εικ. 4.2γ') διακρίνεται από βελτιωμένη αντίθεση, ωστόσο διέπεται και από εμφανή παραμόρφωση και κωδωνισμό. Τέλος, το αποτέλεσμα του αλγορίθμου μας (Εικ. 4.2δ') με παραμέτρους $K \leftarrow$ HEVC και $L \leftarrow$ CxHD, επιδεικνύει βελτιωμένη ποιότητα με ελαττωμένη θόλωση και θόρυβο.

Συνεχίζοντας την αξιολόγησή μας υπό άλλη οπτική, εξετάζουμε τη διακύμανση ποιότητας των ανακατασκευασμένων εικόνων όταν αυξάνεται ο παράγοντας υπερ-δειγματοληψίας f . Χρησιμοποιούμε τη συνάρτηση διασποράς σημείου H της ενότητας 4.2.2 προκειμένου να υπερ-δειγματοληπτήσουμε τις εικόνες αναφοράς *carphone* και *blue sky* (μεγέθους QCIF και 1080p αντίστοιχα) κατά 256 έως 1024 φορές αντίστοιχα ($f \in \{4, 16, 64, 256, 1024\}$). Οι εικόνες αυτές επελέγησαν κα-

Πίνακας 4.1: Μετρική PSNR (dB) για τη Δικυβική παρεμβολή και ΔPSNR για τις υπόλοιπες μεθόδους όταν πραγματοποιείται υπερ-δειγματοληψία κατά $f = 4$.

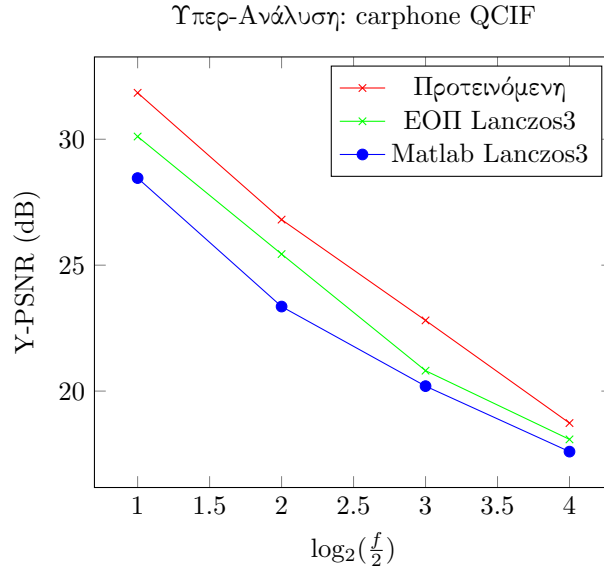
Κ : εικόνα / \mathcal{L} :	Δικυβική Δικυβική		Lcz		AVC		HEVC		CxHD		MLcz	[40]
	εικόνα / \mathcal{L} :	εικόνα / \mathcal{L} :	Lcz	CxHD	Lcz	CxHD	AVC	CxHD	HEVC	CxHD		
carphone QCIF	31.5164	0.1747	0.3467	0.2009	0.3854	0.1028	0.3110	0.1080	0.3268	-0.0237	-3.0583	-1.4089
bridge CIF	27.0208	0.1452	0.1417	0.1318	0.1458	0.1169	0.1341	0.1092	0.1213	-0.1347	-0.7297	-0.2436
bus CIF	28.3510	0.6208	0.5546	0.6543	0.6077	0.6067	0.5638	0.6236	0.6028	-0.9247	-2.4700	-0.9078
coastguard CIF	30.5433	0.3478	0.3847	0.3374	0.3624	0.2829	0.3447	0.2647	0.3391	-0.7161	-2.2498	-1.4428
foreman CIF	32.4661	0.2639	0.3116	0.2430	0.3201	0.3088	0.3561	0.3096	0.3591	0.0737	-1.7323	-0.0663
news CIF	30.9116	0.7892	0.7939	0.7708	0.8164	0.8078	0.8169	0.7969	0.8413	-0.4868	-3.1749	-0.8196
sky SD1	35.8966	1.0113	1.0151	1.0544	1.0185	1.2262	1.2406	1.2805	1.2801	-3.0842	-3.9911	-0.9647
ped. SD1	40.4839	0.2153	0.2874	0.2004	0.2107	0.2919	0.3587	0.3147	0.3592	-1.2490	-3.7159	-1.0394
river SD1	38.0404	0.4440	0.5723	0.6232	0.6447	0.6202	0.7201	0.6511	0.7562	-2.2977	-4.5416	-1.4741
rush SD1	44.8847	0.1646	0.4827	0.2362	0.2632	0.6842	0.8217	0.6964	0.8365	-2.4008	-7.2962	-1.8887
sky 720p	39.1113	0.9852	1.1293	0.8746	0.9127	1.3341	1.4252	1.3880	1.4830	-3.4267	-5.0408	-2.0482
ped. 720p	42.7398	0.2781	0.4456	0.1246	0.0718	0.5101	0.5351	0.5186	0.5351	-1.4390	-4.0709	-1.3790
river 720p	39.5554	0.3334	0.5549	0.4789	0.4816	0.5677	0.7040	0.5947	0.7317	-1.8381	-4.7455	-1.8612
rush 720p	46.3980	-0.1377	0.1595	-0.6618	-0.8063	0.2387	0.2980	0.2357	0.2981	-1.6592	-6.4166	-2.7943
sky 1080p	43.0360	0.7972	1.0810	0.7523	0.7218	1.4551	1.6091	1.5177	1.6706	-3.8388	-6.0555	-3.7413
ped. 1080p	44.4842	-0.0316	0.2343	-0.0933	-0.0234	0.2249	0.3034	0.2264	0.3085	-1.3328	-3.9578	-1.4724
river 1080p	42.1276	0.1439	0.4489	0.3972	0.4254	0.4658	0.6308	0.4859	0.6556	-1.9585	-5.1123	-2.5791
rush 1080p	46.3279	-0.3564	-0.0767	-0.6741	-0.6420	-0.1029	-0.0317	-0.1135	-0.0331	-1.0076	-4.2874	-2.7956
μ.τ.	37.9942	0.3438	0.4926	0.3139	0.3287	0.5412	0.6190	0.5560	0.6373	-1.8553	-4.0359	-1.6070

Πίνακας 4.2: Μετρική MSSIM για τη Δικυβική παρεμβολή και ΔMSSIM για τις υπόλοιπες μεθόδους όταν πραγματοποιείται υπερ-δειγματοληψία κατά $f = 4$.

Κ : εικόνα / \mathcal{L} :	Δικυβική Δικυβική		Lcz		AVC		HEVC		CxHD		MLcz	[40]
	Δικυβική	Δικυβική	Lcz	CxHD	AVC	CxHD	HEVC	CxHD	HEVC	CxHD		
carphone QCIF	0.8942	-0.0024	0.0026	-0.0004	0.0037	-0.0035	0.0025	-0.0029	0.0036	0.0037	-0.0308	-0.0332
bridge CIF	0.8345	0.0027	0.0064	0.0017	0.0036	0.0010	0.0051	0.0007	0.0052	-0.0033	-0.0226	-0.0171
bus CIF	0.8145	0.0163	0.0176	0.0177	0.0184	0.0141	0.0164	0.0145	0.0174	-0.0223	-0.0532	-0.0306
coastguard CIF	0.8066	-0.0149	0.0174	0.0143	0.0157	0.0124	0.0153	0.0120	0.0155	-0.0262	-0.0581	-0.0340
foreman CIF	0.9313	-0.0011	0.0023	-0.0025	-0.0009	0.0003	0.0037	0.0011	0.0043	0.0017	-0.0133	-0.0100
news CIF	0.9213	0.0033	0.0061	0.0036	0.0058	0.0029	0.0059	0.0027	0.0066	-0.0002	-0.0270	-0.0201
sky SD1	0.9650	0.0028	0.0039	0.0006	-0.0001	0.0045	0.0056	0.0050	0.0061	-0.0265	-0.0176	-0.0045
ped. SD1	0.9768	-0.0008	0.0003	-0.0012	-0.0011	0.0002	0.0008	0.0003	0.0009	-0.0084	-0.0094	-0.0042
river SD1	0.9514	0.0016	0.0038	0.0040	0.0039	0.0040	0.0054	0.0043	0.0057	-0.0377	-0.0301	-0.0096
rush SD1	0.9888	-0.0013	-0.0001	-0.0015	-0.0017	0.0005	0.0009	0.0005	0.0010	-0.0095	-0.0081	-0.0018
sky 720p	0.9776	0.0008	0.0021	-0.0029	-0.0039	0.0029	0.0036	0.0032	0.0040	-0.0211	-0.0125	-0.0039
ped. 720p	0.9810	-0.0013	-0.0000	-0.0027	-0.0033	-0.0001	0.0002	-0.0000	0.0003	-0.0064	-0.0072	-0.0036
river 720p	0.9624	-0.0004	0.0021	0.0009	0.0005	0.0020	0.0034	0.0023	0.0037	-0.0248	-0.0228	-0.0085
rush 720p	0.9891	-0.0014	-0.0005	-0.0039	-0.0046	-0.0003	-0.0002	-0.0003	-0.0002	-0.0054	-0.0046	-0.0020
sky 1080p	0.9840	-0.0006	0.0008	-0.0023	-0.0032	0.0017	0.0022	0.0019	0.0024	-0.0153	-0.0073	-0.0032
ped. 1080p	0.9805	-0.0017	-0.0002	-0.0020	-0.0019	-0.0006	-0.0001	-0.0006	-0.0001	-0.0055	-0.0055	-0.0030
river 1080p	0.9720	-0.0014	0.0009	0.0003	0.0003	0.0007	0.0019	0.0009	0.0021	-0.0175	-0.0145	-0.0063
rush 1080p	0.9839	-0.0016	-0.0005	-0.0029	-0.0030	-0.0007	-0.0004	-0.0007	-0.0004	-0.0037	-0.0031	-0.0021
μ.τ.	0.9397	0.0016	0.0036	0.0011	0.0016	0.0023	0.0040	0.0025	0.0043	-0.0127	-0.0193	-0.0110



Εικόνα 4.2: Υποκειμενική σύγκριση αλγορίθμων ΥΑ: 4.2α') εικόνα αναφοράς, 4.2β') δικυβική, 4.2γ') Matlab IBP Lanczos3, 4.2δ') προτεινόμενη.

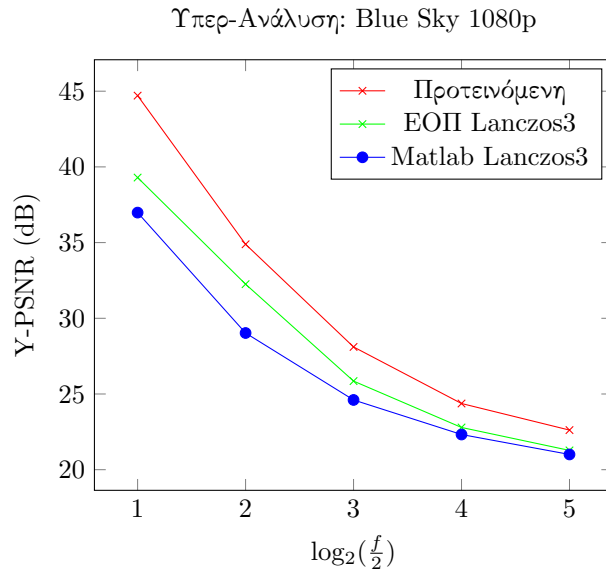


Εικόνα 4.3: Απόδοση ΥΑ (Y-PSNR) συναρτήσει του παράγοντα υπερ-δειγματοληψίας: σύγκριση 3 αλγορίθμων με εικόνα αναφοράς QCIF.

θώς παρείχαν τη χαμηλότερη και υψηλότερη τιμή Δ PSNR κατά τις προηγούμενες μετρήσεις. Τα αποτελέσματα φαίνονται στις καμπύλες 4.3 και 4.4 για 3 αλγορίθμους: *MLez*, *EOΠ [40]*, και τον προτεινόμενο με παραμέτρους $K \leftarrow$ HEVC και $L \leftarrow$ CxHD. Συνολικά, συμπεραίνουμε ότι η ποιότητα ελαττώνεται παράλληλα με την αύξηση του παράγοντα υπερ-δειγματοληψίας και η τιμή του Δ PSNR φαίνεται ισχυρά εξαρτώμενη από την αρχική μέθοδο ανακατασκευής. Επιπρόσθετα, στις περισσότερες των περιπτώσεων η σχετική τιμή του Δ PSNR παρουσιάζει μια συγκριτικά σταθερή συμπεριφορά (με εξαίρεση τον παράγοντα $f = 64$ για τη μέθοδο [40]). Το προτεινόμενο σχήμα υπερτερεί ξεκάθαρα των υπολοίπων με το κέρδος σε PSNR να κυμαίνεται από 0.65 έως 1.74 dB για την εικόνα μεγέθους QCIF και από 1.34 έως 5.41 dB για την εικόνα μεγέθους 1080p (σημειώνουμε ότι σε σύγκριση με την παρεμβολή Lanczos, το κέρδος μπορεί ακόμη και να διπλασιαστεί, όπως φαίνεται από την μπλε καμπύλη).

Κατά την αξιολόγηση και εξερεύνηση του χώρου λύσεων, εκτός από τις παραμέτρους του προτεινόμενου αλγορίθμου ($\{K, L, f\}$ όπως παραπάνω, $\{G, H\}$, όπως στην ενότητα 4.2.2), εστίασαμε και στη βελτίωση του αριθμού των επαναλήψεων n . Ο πίνακας 4.3 συγκεντρώνει το συνολικό αριθμό επαναλήψεων οι οποίες απαιτούνται για τον τερματισμό κάθε σχήματος οπισθοπροβολής. Στις γραμμές 1 και 2 αναφέρεται ο αριθμός επαναλήψεων για το συμβατικό σχήμα EOP [40] με αρχική ανακατασκευή μέσω Δικυβικής ή Lanczos παρεμβολής. Από την τρίτη γραμμή του πίνακα, προκύπτει ότι, στις περισσότερες περιπτώσεις, το προτεινόμενο σχήμα χρειάζεται λιγότερες από 10 επαναλήψεις (και λιγότερες εν συγκρίσει με εκείνες των άλλων δύο αλγορίθμων).

Για να αξιολογήσουμε την πολυπλοκότητα του αλγορίθμου μας, καταμετρούμε



Εικόνα 4.4: Απόδοση ΥΑ (Υ-PSNR) συναρτήσει του παράγοντα υπερ-δειγματοληψίας: σύγκριση 3 αλγορίθμων με εικόνα αναφοράς 1080p.

Πίνακας 4.3: Απαιτούμενες επαναλήψεις για τη σύγκλιση σχημάτων EOI.

	carphone	coastguard	sky SD1	ped. SD1	river SD1	rush SD1	sky 720p	ped. 720p	river 720p	rush 720p	sky 1080p	ped. 1080p	river 1080p	rush 1080p
IBP-Δικυβική	23	17	9	14	11	12	7	11	8	8	4	8	6	4
IBP-Lanczos	23	18	10	15	11	12	7	11	9	8	4	9	6	4
HEVC-CxHD	8	10	10	8	9	8	9	9	9	8	9	9	9	7

το χρόνο εκτέλεσης σε έναν 8-πύρηνο x86-64 επεξεργαστή γενικού σκοπού, χρονισμένου στα 3.6GHz. Η ανακατασκευή ($f = 4$) μίας εικόνας 1080p, απαιτεί κατά μέσο όρο 3.3 sec με 19% απόκλιση, ανάλογα με τη διαμόρφωση του σχήματος (η διαμόρφωση *Δικυβική-Δικυβική* είναι η ταχύτερη, ενώ η διαμόρφωση *HEVC-CxHD* απαιτεί τον περισσότερο χρόνο). Σε εικόνες μεγέθους QCIF, ο ανωτέρω χρόνος περιορίζεται σε 0.05 sec. Από την αποτίμηση της κατανομής του χρόνου εκτέλεσης του αλγορίθμου (execution profiling), βλέπουμε ότι περίπου το 57% του χρόνου δαπανάται κατά τη συνέλιξη με *G* και *H*, ενώ το 32% κατά την υπερ-δειγματοληψία *K* και *L*. Σε σύγκριση με τις απλές μεθόδους παρεμβολής ο προτεινόμενος αλγόριθμος απαιτεί όπως αναμενόταν έως και δύο τάξεις μεγέθους περισσότερο χρόνο. Από την άλλη πλευρά, η αντιπαραβολή με άλλες μεθόδους ΥΑ της ίδιας κατηγορίας, καταδεικνύει επιτάχυνση της εκτέλεσης έως και κατά 5 φορές (υλοποίηση του σχήματος *HEVC-CxHD* σε Matlab 2011b και σύγκριση με τη μέθοδο [40] σε εικόνες αναφοράς 1080p). Κατά συνέπεια, επιτυγχάνεται σημαντική βελτίωση αποδοτικότητας, δεδομένης και της ενίσχυσης ποιότητας η οποία καθίσταται εφικτή με μειωμένο υπολογιστικό φόρτο.

ΚΕΦΑΛΑΙΟ 5

Μειωμένης Πολυπλοκότητας ΥΑ για Συμπίεση Εικονορροών Χαμηλού Δυφιορρυθμού

Βασιζόμενοι στα ευρήματα των προηγούμενων κεφαλαίων, επεκτείνουμε την εξερεύνησή μας με διττό στόχο: α) αφενός, να μειώσουμε την πολυπλοκότητα ενός συστήματος κωδικοποίησης/αποκωδικοποίησης εικονορροών χαμηλού δυφιορρυθμού και αφετέρου, β) να βελτιώσουμε παράλληλα την προκύπτουσα ποιότητα. Για το σκοπό αυτό, παρουσιάζουμε μία λύση η οποία επιτυγχάνει υψηλή απόδοση δυφιορρυθμού-παραμόρφωσης και βελτιωμένες επιδόσεις για δυφιορρυθμούς έως 3.8 Mbps και 10 Mbps στα πρότυπα H.264 και HEVC αντιστοίχως. Η συμβολή μας έγκειται στη βελτίωση ενός γενικευμένου σχήματος αποδεκότητας/χβάντισης μελετώντας το ως ένα τρισδιάστατο πρόβλημα βελτιστοποίησης. Αποσκοπούμε στη βελτίωση του τρίπτυχου δυφιορρυθμού - παραμόρφωσης - πολυπλοκότητας και όχι μόνο στο δίπτυχο δυφιορρυθμού - παραμόρφωσης, όπως είναι η συνηθισμένη προσέγγιση στη βιβλιογραφία. Με τη μελέτη μας προάγουμε το ανωτέρω σχήμα συμπίεσης, βελτιστοποιώντας και εξερευνώντας από κοινού τις ανεξάρτητες διαδικασίες αποδεκότητας-υπερ-δειγματοληψίας και προτείνοντας τεχνικές χαμηλού κόστους για τη διεκπεραίωσή τους. Εξ όσων γνωρίζουμε, η παρούσα μελέτη είναι η πρώτη στη βιβλιογραφία η οποία εφαρμόζει και μελετά ένα τέτοιο σχήμα συμπίεσης σε εικόνες υπερυψηλής ανάλυσης (Ultra-High Definition - UHD) μέσω του προτύπου HEVC. Το πιο σημαντικό τμήμα της μελέτης αυτής, επιπρόσθετα από τη βελτίωση του σχήματος κωδικοποίησης, είναι η πρόταση ενός νέου αλγορίθμου υπερ-δειγματοληψίας. Ο προτεινόμενος αλγόριθμος συνδυάζει υψηλή ποιότητα αποτελεσμάτων με μειωμένη πολυπλοκότητα και μπορεί να χρησιμοποιηθεί ως αυτόνομη λύση ΥΑ εικόνων.

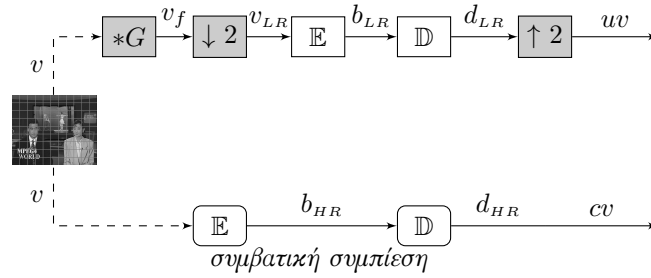
Ειδικά στο πεδίο της ΥΑ, η μελέτη μας περιλαμβάνει την έρευνα για την ανάπτυξη μιας λύσης ΥΑ η οποία εξυπηρετεί το σκοπό της τρισδιάστατης βελτιστοποίησης δυφιορρυθμού - παραμόρφωσης - πολυπλοκότητας. Εισαγάγουμε λοιπόν μία νέα τεχνική ΥΑ η οποία βασίζεται στο πλαίσιο ΕΟΠ (*IBP*). Χαρακτηρίζεται από: α) είσοδο μίας εικόνας, β) υπολογισμό ανακατασκευής μέσω χαμηλής πολυπλοκότητας φίλτρων 2 και 4 δειγμάτων, γ) ανίχνευση ακμών μέσω κλίσης και δ) συνδυασμό των ανωτέρω χαρακτηριστικών σε μία ακμοστρεφή φάση ανακατα-

σκευής και μία προσαρμοστική φάση βελτίωσης σφάλματος η οποία συγκλίνει σε περιορισμένο αριθμό επαναλήψεων. Πέρα από το πλαίσιο της συμπίεσης εικονορορών, η παρούσα μελέτη αξιολογεί τις επιδόσεις της προτεινόμενης τεχνικής ως αυτόνομης λύσης ΥΑ γενικής χρήσης. Ο προτεινόμενος αλγόριθμος επιτυγχάνει: α) ανώτερη ποιότητα αποτελεσμάτων κατά 5 dB σε σχέση με το απλό φιλτράρισμα και έως 2 dB σε σχέση με άλλα σχήματα ΕΟΠ, β) συγκρίσιμες τιμές PSNR και παραπλήσιες τιμές MSSIM σε σχέση με *λεξικοπαγείς/αραιής αναπαράστασης* μεθόδους οι οποίες αντιπροσωπεύουν την τελευταία λέξη της τεχνολογίας, γ) την υψηλότερη BRISQUE μετρική από όλες τις μεθόδους εισόδου μίας εικόνας και δ) χαμηλότερο χρόνο υπολογισμού, μικρότερο κατά τρεις τάξεις μεγέθους εν συγκρίσει με τις πιο εκλεπτυσμένες υπερσύγχρονες μεθόδους ΥΑ.

Προκειμένου να βελτιώσουμε τη συνολική πολυπλοκότητα του συστήματος, η μελέτη μας συνδυάζει δύο προσεγγίσεις. Πρώτον, θεωρούμε ότι ο κωδικοποιητής επεξεργάζεται πλαίσια χαμηλότερης ανάλυσης: αποδεκατίζοντας την είσοδο κατά παράγοντα 2 (η βέλτιστη τιμή κατά τη σχετική βιβλιογραφία) σχεδόν υποτετραπλασιάζουμε τον υπολογιστικό φόρτο του κωδικοποιητή. Δεδομένου ότι ο τελευταίος συνιστά και το πιο απαιτητικό υπολογιστικά τμήμα, εν δυνάμει υποτετραπλασιάζουμε και την πολυπλοκότητα όλου του συστήματος κωδικοποίησης/αποκωδικοποίησης. Αυτό αποτελεί σημαντική ωφέλεια δεδομένου του μεγάλου αριθμού εικονοστοιχείων σε εικόνες υψηλής και υπερ-υψηλής ανάλυσης. Δεύτερον, δίνουμε ιδιαίτερη βαρύτητα στην επιλογή του αλγορίθμου υπερ-δειγματοληψίας στην πλευρά του αποκωδικοποιητή. Εξερευνούμε το πεδίο της ΥΑ αναπτύσσοντας και συγκρίνοντας τις επιδόσεις ποιότητας και το χρόνο εκτέλεσης αντιπροσωπευτικών τεχνικών ΥΑ. Τέλος, προσδιορίζουμε τα αποτελέσματά τους όταν εφαρμοστούν στο βελτιωμένο σχήμα συμπίεσης και προτείνουμε ένα νέο αλγόριθμο ο οποίος βελτιώνει τις επιδόσεις του σχήματος έχοντας χαμηλό υπολογιστικό κόστος.

5.1 Προτεινόμενο σχήμα συμπίεσης

Στην παρούσα ενότητα ακολουθούμε το σχήμα αποδεκάτισης/κβάντισης της εικόνας 5.1. Εν συγκρίσει με έναν συμβατικό κωδικοποιητή, το σχήμα περιλαμβάνει μία διάκριτη μονάδα αποδεκάτισης στην πλευρά του κωδικοποιητή και μία διάκριτη μονάδα υπερ-δειγματοληψίας στην πλευρά του αποκωδικοποιητή. Η μονάδα αποδεκάτισης παράγει τις εικονορορές χαμηλής ανάλυσης v_{LR} μέσω μίας υπομονάδας συνέλιξης ($*G$) με πυρήνα G , και ενός απλού αποδεκατιστή κατά παράγοντα 2 (δηλαδή τελικώς διατηρείται το 1/4 των εικονοστοιχείων). Οι μονάδες κωδικοποίησης και αποκωδικοποίησης έχουν συμβατική λειτουργία, γεννώντας αντίστοιχα το κωδικοποιημένο δυφίορρευμα (coded bitstream) b_{LR} , και τα αποκωδικοποιημένα πλαίσια, d_{LR} . Η μονάδα υπερ-δειγματοληψίας περιλαμβάνει τον αλγόριθμο ΥΑ από τον οποίο δημιουργείται η εικονορορή αρχικού μεγέθους (uv), μέσω υπερ-δειγματοληψίας των αποκωδικοποιημένων πλαισίων κατά παράγοντα 2 (σε κάθε διεύθυνση). Εφεξής λοιπόν, εστιάζουμε στη διερεύνηση/σχεδίαση αλγορίθμων για κάθε μία από τις τέσσερις προαναφερθείσες μονάδες, και ειδικότερα



Εικόνα 5.1: Συμβατικό και προτεινόμενο σχήμα κωδικοποίησης.

για τη μονάδα υπερ-δειγματοληψίας, στοχεύοντας στη βελτίωση της ποιότητας της εικονορροής uv με παράλληλη μείωση του μεγέθους του δυοιορρέυματος b_{LR} .

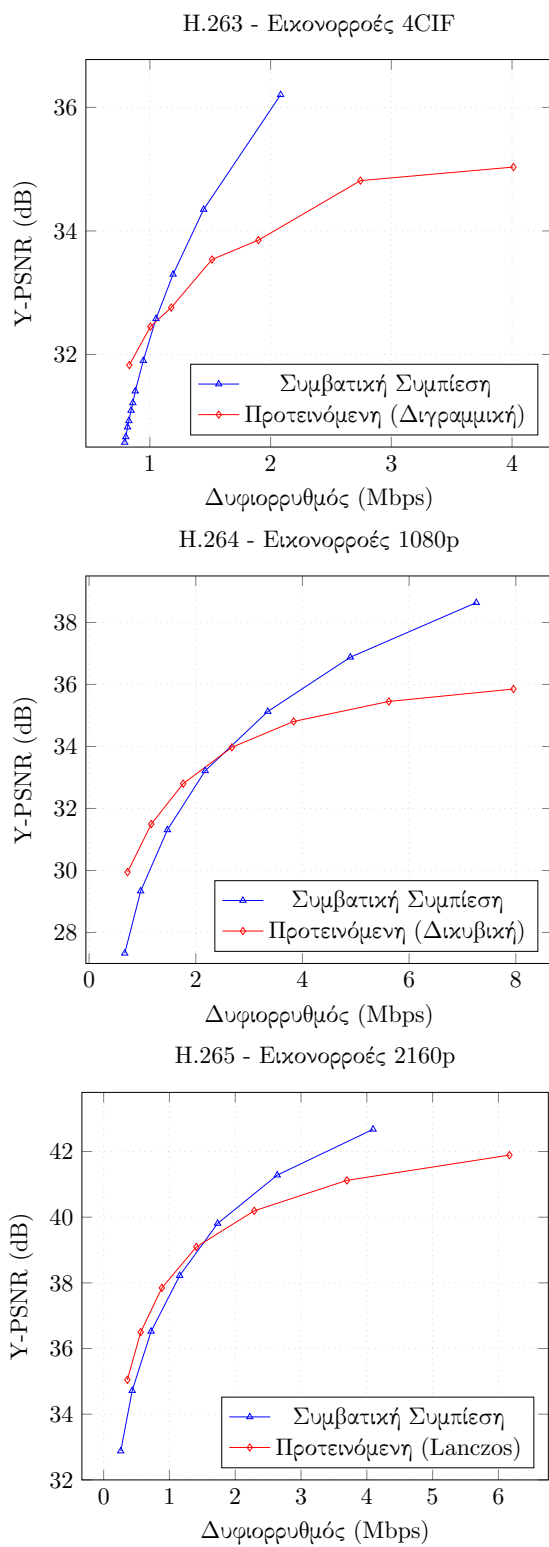
Ξεκινούμε τη μελέτη μας από τις μονάδες κωδικοποίησης και αποκωδικοποίησης, οι οποίες οφείλουν να ακολουθούν ένα από τα συνήθη πρότυπα συμπίεσης. Εξετάζουμε τρία χαρακτηριστικά πρότυπα συμπίεσης εικονορροών τα οποία αντιπροσωπεύουν τρεις διαδοχικές γενιές τεχνολογίας κωδικοποίησης: το H.263, το H.264, και το πιο πρόσφατο HEVC (ή αλλιώς H.265). Η εξέταση τριών διαδοχικών γενεών κωδικοποίησης έχει διπλό όφελος: πρώτον, αποδεικνύουμε ότι το προτεινόμενο σχήμα συμπίεσης υπερισχύει σε χαμηλούς δυοιορρυθμούς έναντι της συμβατικής κωδικοποίησης, παρά τη διαρκή εξέλιξη της τελευταίας και την ελάττωση του θορύβου κβάντισης την οποία αυτή συνεπάγεται. Δεύτερον, με τον τρόπο αυτό καθίσταται δυνατόν να παρουσιάσουμε την επίδραση διάκριτων μονάδων συμπίεσης στις επιδόσεις και την πολυπλοκότητα του προτεινόμενου σχήματος.

Χωρίς σημαντική βλάβη της γενικότητας και συνυπολογίζοντας το στόχο μας να διατηρήσουμε την πολυπλοκότητα σε χαμηλά επίπεδα, ρυθμίζουμε τους συμβατικούς κωδικοποιητές για μέσο- έως ανώτερο- φόρτο (π.χ., το *baseline* προφίλ για το H.264 και το *main 12* για το H.265). Συν τοις άλλοις, επιλέγουμε χάριν παραδείγματος, μία γκαουσιανή ΣΔΣ μεγέθους 5×5 για τη μονάδα αποδεκάτισης και απλές μεθόδους παρεμβολής για τη μονάδα υπερ-δειγματοληψίας: δύο δειγμάτων *Διγραμμική*, 4 δειγμάτων *Δικυβική* και 6 δειγμάτων *Lanczos* (όλοι οι συντελεστές αντιστοιχούν σε διαχωρίσιμα φίλτρα). Κατά τις δοκιμές μας εκτελούμε μονονηματικές εφαρμογές σε έναν επεξεργαστή γενικού σκοπού (αρχιτεκτονικής x86-64 τεσσάρων πυρήνων, χρονισμένο στα 2.8 GHz, με 4 GB μνήμη συστήματος). Οι αλγόριθμοι μετασχηματισμού μεγέθους υλοποιήθηκαν στα πλαίσια της διατριβής σε γλώσσα C, ενώ για τους κωδικοποιητές/αποκωδικοποιητές βασισόμαστε στο μοντέλο λογισμικού αναφοράς [1], [5], δηλαδή, το TMN 2.0 για το H.263, το JM 17 για το H.264, και το HM 9.2rc1 για το H.265. Η είσοδος στο σχήμα προέρχεται από τον ακόλουθο συνδυασμό {Εικονορροών}/{Αναλύσεων} (με ρυθμό 25 πλαίσια ανά sec): {Carphone}/{QCIF}, {Bridge, Bus, Coastguard, Foreman, News}/{CIF}, {Blue Sky, Pedestrian, Riverbed, Rush Hour}/{4CIF, SD 720×576, HD 720p, HD 1080p}, {Shields}/{HD 720p} και {Cactus, Coast-

guard, Foreman, Mobile, News, Park, Suzie}/ {UHD 2160p}. Για κάθε ανάλυση και σχήμα συμπίεσης (συμβατικό ή προτεινόμενο) υπολογίζουμε το μέσο δυφιορρυθμό και τη μέση τιμή της μετρικής PSNR όλων των εικονορροών, προκειμένου να σχεδιάσουμε τις αντίστοιχες καμπύλες δυφιορρυθμού - παραμόρφωσης.

Στην εικόνα 5.2 παρουσιάζονται τρία αντιπροσωπευτικά ζεύγη καμπυλών: ένα για το H.263 με *Διγραμμική* παρεμβολή (σε 4CIF ανάλυση, ήτοι 704×576 εικονοστοιχεία), ένα για το H.264 με *Δικυβική* παρεμβολή (σε 1080p ανάλυση, ήτοι 1920×1080 εικονοστοιχεία) και ένα για το H.265 με παρεμβολή *Lanczos* (σε ανάλυση UHD). Για κάθε ζεύγος, συγκρίνουμε την απόδοση του συμβατικού κωδικοαποκωδικοποιητή έναντι αυτής του προτεινόμενου σχήματος με τις ίδιες μονάδες κωδικοποίησης/αποκωδικοποίησης. Σε όλες τις περιπτώσεις, εμφανίζεται ένα κατώφλι “κρίσιμου δυφιορρυθμού” (επικάλυψη καμπυλών), κάτω από το οποίο το προτεινόμενο σχήμα υπερσχύει έναντι της συμβατικής συμπίεσης. Η τιμή του κρίσιμου δυφιορρυθμού και το ακριβές κέρδος σε PSNR εξαρτώνται από την ανάλυση των εικονορροών, από το πρότυπο κωδικοαποκωδικοποίησης και προφανώς, τον αλγόριθμο υπερ-δειγματοληψίας. Όπως φαίνεται στην εικόνα 5.2, όταν κωδικοποιούμε είσοδο μέσης ανάλυσης με το H.263 ο κρίσιμος δυφιορρυθμός είναι 1 Mbps και το κέρδος σε PSNR είναι μικρότερο των 0.8 dB. Ωστόσο, όταν επεξεργαζόμαστε εικονορροές υψηλής ανάλυσης με κάποιο από τα σύγχρονα πρότυπα κωδικοποίησης (δοκιμές σε H.264 και H.265), ο κρίσιμος δυφιορρυθμός αυξάνεται στην περιοχή των 3 Mbps ενώ το μέσο κέρδος σε PSNR μπορεί να φτάσει και τα 2.4 dB. Να σημειωθεί εδώ ότι η τιμή των 3 Mbps αποτελεί ένα ρεαλιστικό εύρος ζώνης για πολλές βιομηχανικές εφαρμογές (για παράδειγμα, την ψηφιακή μετάδοση τηλεοπτικού σήματος).

Από περαιτέρω ανάλυση των καμπυλών συμπεραίνουμε ότι ο κρίσιμος δυφιορρυθμός αυξάνεται ομαλά με την αύξηση του μεγέθους των πλαισίων εισόδου. Συγκεκριμένα στην περίπτωση του H.264, ο κρίσιμος δυφιορρυθμός αυξάνεται από 0.4 Mbps σε SD εικονορροές, σε 0.95 Mbps σε 720p και 2.5 Mbps σε 1080p πλαίσια. Σε χαμηλούς δυφιορρυθμούς, το προτεινόμενο σχήμα υπερσχύει του συμβατικού σε ποιότητα κατά 1.34 dB για SD πλαίσια έως 2.4 dB για 1080p πλαίσια (PSNR μετρική). Κατά τις δοκιμές με το H.265 και όταν χρησιμοποιούμε *Δικυβική* παρεμβολή (ίδια με την περίπτωση του H.264), ο κρίσιμος δυφιορρυθμός αυξάνεται με παρόμοιο τρόπο έως και τα 0.22 Mbps σε 1080p πλαίσια. Η τιμή αυτή είναι σημαντικά μικρότερη εκείνης με το H.264, λόγω της εξαιρετικής αποδοτικότητας κωδικοποίησης του HEVC προτύπου. Εξετάζοντας απομονωμένα κάποιον κωδικοαποκωδικοποιητή σε συγκεκριμένη ανάλυση εισόδου, παρατηρούμε μία αξιοσημείωτη διακύμανση τόσο στον κρίσιμο δυφιορρυθμό, όσο και στο Δ PSNR αναλόγως του περιεχομένου της εικονορροής. Για παράδειγμα, σε 1080p πλαίσια επεξεργασμένα με το H.264 ο κρίσιμος δυφιορρυθμός κυμαίνεται από τα 0.9 Mbps (εικονορροή *BlueSky*) έως τα 8.1 Mbps (*Riverbed*). Επίσης, το μέγιστο κέρδος ποιότητας κυμαίνεται από 1 (*Riverbed*) έως 5 dB (*BlueSky*). Όταν επεξεργαζόμαστε πλαίσια μεγέθους 2160p με το H.265 και *Lanczos* παρεμβολή, ο κρίσιμος δυφιορρυθμός κυμαίνεται από 0.25 έως 12.68 Mbps και το κέρδος ποιότητας του προτεινόμενου σχήματος βρίσκεται στο διάστημα 0.57 (εικονορροή *Park*) έως



Εικόνα 5.2: Επιδόσεις του προτεινόμενου σχήματος συμπίεσης με τη χρήση διακριτών προτύπων κωδικοποίησης και φίτρων παρεμβολής (25fps).

2.45 dB (εικονορροή *Cactus*).

Προκειμένου να αποτιμήσουμε την πολυπλοκότητα των σχημάτων, αναλύουμε το χρόνο εκτέλεσης κατά την επεξεργασία πλαισίων μεγέθους 1080p. Από την ανάλυση προκύπτει ότι ο συμβατικός κωδικοποιητής του H.264 απαιτεί περίπου 14 δευτερόλεπτα για τη συμπίεση ενός πλαισίου 1920 × 1080 εικονοστοιχείων, ενώ ο αποκωδικοποιητής ολοκληρώνει την εργασία του σε 0.08 sec. Η συμβατική συμπίεση με το H.265 σχεδόν πενταπλασιάζει το χρόνο κωδικοποίησης (περίπου 65 sec) και ο αποκωδικοποιητής του προτύπου απαιτεί 0.1 δευτερόλεπτα για το ίδιο μέγεθος πλαισίου. Στον αντίποδα, το προτεινόμενο σχήμα ελαττώνει το χρόνο κωδικοποίησης του H.264 σε 4 δευτερόλεπτα ανά πλαίσιο, προσθέτοντας μόλις 0.1 sec για τη διαδικασία της αποδεκμάτισης. Αυτό συνεπάγεται σχεδόν 3.5x επιτάχυνση στην πλευρά του κωδικοποιητή. Στην πλευρά του αποκωδικοποιητή, το πρότυπο H.264 απαιτεί 0.026 sec και η Δικυβική παρεμβολή προσθέτει στο χρόνο αυτό 0.021 δευτερόλεπτα ανά πλαίσιο. Συνεπώς, επιταχύνουμε (σχεδόν κατά 2x) και την πλευρά του αποκωδικοποιητή. Παρόμοιες τιμές επιτάχυνσης επιτυγχάνονται και στην περίπτωση του προτύπου H.265.

Η ανωτέρω ανάλυση μας δείχνει ότι με το προτεινόμενο σχήμα μπορούμε να ελαττώσουμε την πολυπλοκότητα, ειδικά στην πλευρά του κωδικοποιητή και ταυτόχρονα, να επιτύχουμε βελτιωμένη ποιότητα σε διάφορες εφαρμογές συμπίεσης. Αναλόγως των απαιτήσεων της εφαρμογής, μπορούμε να χρησιμοποιήσουμε απλές τεχνικές υπερ-δειγματοληψίας (π.χ. παρεμβολή με φίλτρα ακεραίων συντελεστών 2 έως 6 δειγμάτων) προκειμένου να ελαττώσουμε το χρόνο αποκωδικοποίησης σε σχέση με τη συμβατική διαδικασία των H.264/H.265, ή, να βελτιώσουμε περαιτέρω την ποιότητα της προκύπτουσας εικονορροής μέσω πιο προηγμένων τεχνικών όπως αυτές που θα παρουσιάσουμε στην αμέσως επόμενη ενότητα.

5.2 Μελέτη και ανάπτυξη αλγορίθμων ΥΑ

Η τρέχουσα αλλά και οι ακόλουθες ενότητες, εστιάζουν στη βελτίωση της απόδοσης του προτεινόμενου σχήματος συμπίεσης μέσω προηγμένων αλγορίθμων ΥΑ στην πλευρά του αποκωδικοποιητή. Για το σκοπό αυτό, ξεκινούμε τη μελέτη μας από την εξέταση τεχνικών ΥΑ ως αυτόνομων λύσεων (εκτός του πλαισίου συμπίεσης). Στην τρέχουσα ενότητα, αποτιμούμε ένα αντιπροσωπευτικό σύνολο από πολύ γνωστούς και ευρέως χρησιμοποιημένους αλλά και υπερσύγχρονους αλγορίθμους ΥΑ. Επιπρόσθετα, παρουσιάζουμε έναν νέο αλγόριθμο ΥΑ μέσω του οποίου σκοπεύουμε να παράσχουμε ποιότητα εικόνας αντίστοιχη των υπερσύγχρονων μεθόδων και παράλληλα να διατηρήσουμε σχετικά χαμηλά το υπολογιστικό κόστος.

Το σύνολο των εξεταζομένων αλγορίθμων περιλαμβάνει απλές μεθόδους υπερ-δειγματοληψίας, μεθόδους *EOPI*, καθώς και υπερσύγχρονες *λεξικοπαγείς* μεθόδους ΥΑ βασισμένες σε *αραιή αναπαράσταση*.

Με τον όρο *απλή* υπερ-δειγματοληψία αναφερόμαστε σε συνηθισμένες τεχνικές

παρεμβολής, όπως τη *Διγραμμική*, τη *Δικυβική* ή τη *Lanczos*. Στην πράξη, όλες οι τεχνικές αυτές εκτελούν συνέλιξη του διάκριτου σήματος εισόδου με έναν δισδιάστατο πυρήνα συνέλιξης μεγέθους ίσου με αυτού των δειγμάτων της εκάστοτε μεθόδου. Σε όλες τις περιπτώσεις παράγουμε εκ των προτέρων τους σταθερούς συντελεστές των πυρήνων επιλύοντας εξισώσεις τμηματικών ή μη-πολυωνύμων. Το υπολογιστικό κόστος των *απλών* μεθόδων είναι προκαθορισμένο και διαμορφώνεται κυρίως από τη δισδιάστατη διαδικασία συνέλιξης: περισσότερα σημεία συνεπάγονται υψηλότερο κόστος και βελτιωμένη ποιότητα. Για μια πιο αναλυτική αποτίμηση της απόδοσης των μεθόδων αυτών, ο αναγνώστης μπορεί να ανατρέξει στο [55], και τις ενότητες [29] [30]. Από εδώ και στο εξής, επιλέγουμε μία αντιπροσωπευτική απλή τεχνική την οποία χρησιμοποιούμε ως σημείο αναφοράς έναντι των προηγμένων μεθόδων: τη *Δικυβική* παρεμβολή του MATLAB R2012b. Ο λόγος για τον οποίο επιλέγουμε το φίλτρο αυτό, είναι επειδή αποτελεί την πιο συνηθισμένη μέθοδο στη σχετική βιβλιογραφία (για δίκαια πειραματικά αποτελέσματα) [27] [72] [22] [89] [98].

Το πλαίσιο επεξεργασίας *ΕΟΠ* βασίζεται σε ένα σύνολο από n εικόνες παρατήρησης XE , O_{LR}^n . Το σύνολο των εικόνων αυτών υφίσταται επεξεργασία μέσω συνέλιξης με μία συνάρτηση παραθύρου η οποία επαγάγει κυκλική συμμετρία επί των εικονοστοιχείων. Με το βήμα αυτό βελτιώνεται η ακρίβεια της επακόλουθης διαδικασίας *καταχώρισης εικόνων* (image registration), στην οποία πραγματοποιείται εκτίμηση μετατόπισης και στροφής εικονοστοιχείων. Κατόπιν, επεξεργαζόμαστε την εικόνα επαναληπτικά μέσω ενός βρόχου ο οποίος περιλαμβάνει τα εξής βήματα: α) αντιστάθμιση μετατόπισης και στροφής βάσει των τιμών που εκτιμήθηκαν, β) θόλωση, αποδεκάτιση κατά παράγοντα f και αφαίρεση του αποτελέσματος από την αρχική παρατήρηση, γ) αντιστροφή των προηγούμενων φαινομένων μέσω υπερ-δειγματοληψίας, φιλτραρίσματος και ανεστραμμένης μετατόπισης/στροφής, και τελικώς, δ) κανονικοποίηση μέσω μιας προκαθορισμένης ή προσαρμοστικά υπολογιζόμενης παραμέτρου σφάλματος. Η μελέτη μας περιλαμβάνει την αποτίμηση της συμβατικής *ΕΟΠ* [40] και της μη-τοπικής *ΕΟΠ* (MTEOP, Nonlocal IBP - *NLIBP*) [22]. Επιπροσθέτως, επιλέγουμε το γενικευμένο πλαίσιο *ΕΟΠ* ως βάση του προτεινόμενου αλγορίθμου ΥΑ (Ενότητα 5.2.1). Ο προτεινόμενος αλγόριθμος προάγει το σχήμα *ΕΟΠ* βελτιώνοντας την προσαρμοστικότητα σε κάθε βήμα της διαδικασίας και χρησιμοποιώντας υπολογισμούς χαμηλής πολυπλοκότητας. Συνοπτικά, χρησιμοποιεί με προσαρμοστικό τρόπο *απλές* τεχνικές παρεμβολής και εκτελεί ακοπροσαρμοστική *ΕΟΠ* σε περιπτώσεις εισόδου μίας εικόνας.

Για να κατανοήσουμε τη θέση των παραπάνω *απλών* τεχνικών, συγκρίνουμε την απόδοσή τους με εκείνη υπερσύγχρονων αλγορίθμων εξαιρετικά υψηλής ποιότητας και πολυπλοκότητας. Συγκεκριμένα, εξετάζουμε: α) την ΥΑ *αραιής αναπαράστασης* συνδυασμένη με *λεξικοπαγή* προσέγγιση [98], [23], [86], β) την ΥΑ *πολλαπλών εικόνων* εισόδου με *Bayesian πρότερη γνώση* [90], [89]. Όσον αφορά τις μεθόδους [98], [23], [86], η επεξεργασία πραγματοποιείται σε δύο στάδια: α) κατάρτιση λεξικού και β) ανακατασκευή ΥΑ. Η κατάρτιση του λεξικού στο [98] ξεκινά με τυχαία δειγματοληψία πλακιδίων εικόνας από ένα βασικό σύνολο εικό-

νων με προκαθορισμένες παραμέτρους. Η μέθοδος αυτή χρησιμοποιεί Δικυβική παρεμβολή προκειμένου να παράξει μία αρχική προσέγγιση ΧΕ, από την οποία θα εξαχθούν τα πλακίδια ΧΕ μέσω μητρών κλίσης πρώτης και δεύτερης τάξης. Εν συνεχεία, τα πλακίδια για τα οποία η τιμή της μεταβλητότητας είναι μικρότερη από 10 αφαιρούνται από το τελικό σύνολο. Τελικά, χρησιμοποιώντας συντελεστή κανονικοποίησης $\lambda = 0.15$ για την αραιή αναπαράσταση, τα πλακίδια κανονικοποιούνται και καταρτίζονται από κοινού τα λεξικά ΧΕ και ΥΕ. Η διαδικασία κατάρτισης περιλαμβάνει τον υπολογισμό ενός μοντέλου αραιής αναπαράστασης και την εφαρμογή αυτού για την εξαγωγή του καταρτισμένου λεξικού. Για τον υπολογισμό του μοντέλου αραιής αναπαράστασης χρειάζεται η βελτιστοποίηση ενός γραμμικά περιορισμένου προβλήματος με ορθογωνική αντικειμενική συνάρτηση μέσω της συζυγούς κλίσης για μέγιστο αριθμό 40 επαναλήψεων. Κατά το στάδιο ανακατασκευής, η εικόνα ΧΕ αρχικά υπερ-δειγματοληπτείται μέσω Δικυβικής παρεμβολής και τα χαρακτηριστικά της εξάγονται από συναρτήσεις κλίσης πρώτης και δεύτερης τάξης. Εν συνεχεία, για κάθε πλακίδιο ΧΕ, οι παράμετροι του μοντέλου αραιής ανάκτησης υπολογίζονται από την ίδια μέθοδο ορθογώνιου προγραμματισμού χρησιμοποιώντας το λεξικό ΧΕ.

Όσον αφορά τη φάση ανακατασκευής της μεθόδου [98], θέτουμε το βάρος κανονικοποίησης της αραιής αναπαράστασης σε 0.2, το πλάτος επικάλυψης πλακιδίων σε 4 και χρησιμοποιούμε λεξικό 1024 ζευγών ατόμων διαστάσεων 5×5 εικονοστοιχείων. Προκειμένου να ελέγξουμε τη μέθοδο στο [23], θέτουμε τη μεταβλητότητα του προσθετικού γκαουσιανού θορύβου στο 0 και χρησιμοποιούμε το λεξικό από το πρώτο σύνολο εικόνων για την ανακατασκευή. Σε πιο πρόσφατες έρευνες, οι συγγραφείς της [86] πρότειναν τον υπολογισμό πλησιέστερων γειτόνων καθώς και μίας μήτρας προβολής στο χώρο ΥΕ για κάθε άτομο του χρησιμοποιούμενου λεξικού, με συνέπεια μειωμένο χρόνο υπολογισμού της ανακατασκευασμένης εικόνας. Για την εξέταση της μεθόδου [86] χρησιμοποιήσαμε ένα λεξικό 1024 ατόμων. Για την εξέταση των Bayesian μεθόδων στα [90], [89], χρησιμοποιήσαμε 8 εικόνες εισόδου στις οποίες εφαρμόσαμε τυχαία στροφή/μετατόπιση καθώς και μία γκαουσιανή $\Sigma\Delta\Sigma$ ($\sigma=1.0$).

5.2.1 Προτεινόμενος αλγόριθμος ΥΑ: L-SEABI

Σε αυτή την ενότητα, εισαγάγουμε τον αλγόριθμο χαμηλής πολυπλοκότητας, στατιστικά ακμοπροσαρμοστικής, οπισθοπροβαλλόμενης παρεμβολής, (*Low-Complexity Statistical Edge-Adaptive Back-projected Interpolation - L-SEABI*). Ο αλγόριθμος αυτός σχεδιάστηκε με γνώμονα τη βελτίωση ποιότητας και τη διατήρηση του υπολογιστικού κόστους σε σημαντικά χαμηλότερα επίπεδα εν συγκρίσει με τις μεθόδους τεχνολογικής αιχμής. Συνίσταται σε δύο κύριες φάσεις επεξεργασίας, α) τη φάση δόμησης, η οποία εκτελείται μία φορά και β) τη φάση βελτίωσης, η οποία εκτελείται κατ' επανάληψη. Κατά τη φάση δόμησης, ο αλγόριθμος υπερ-δειγματοληπτεί την εικόνα εισόδου προσαρμοζόμενος στο περιεχόμενο αυτής τοπικά αλλά και σε στατιστική πληροφορία από το σύνολο της εικόνας. Κατά τη διάρκεια της δόμησης, ανιχνεύουμε αιχμές και συγκεντρώνουμε

πληροφορία τόσο για τα εικονοστοιχεία εισόδου, όσο και για τα ζητούμενα. Η πληροφορία αυτή χρησιμοποιείται ως καθοδήγηση για α) τον υπολογισμό της μήτρας ΥΕ στο τέλος της φάσης δόμησης, β) το σύνολο της ακόλουθης διαδικασίας βελτίωσης. Εκτός της προσαρμοστικότητας, μία από τις βασικές ιδέες πίσω από τον αλγόριθμο *L-SEABI* είναι η εκούσια αύξηση της πολυπλοκότητας στη φάση δόμησης - και κατά συνέπεια της απόδοσης του αλγορίθμου - προκειμένου να συγκλίνει ταχύτερα η φάση βελτίωσης και να μειωθεί ο συνολικός χρόνος υπολογισμού. Μία ακόμη βασική ιδέα γύρω από τον αλγόριθμο *L-SEABI* είναι η επιτάχυνση των υπολογισμών αποφεύγοντας τους συντελεστές κινήτης υποδιαστολής και εκμεταλλευόμενοι τα προτερήματα που παρέχουν οι υπολογισμοί και η αποθήκευση ακεραίων (με άλλα λόγια, αποθηκεύουμε τους συντελεστές ως ακεραίους και υλοποιούμε πολλαπλασιασμούς/διαίρεσεις μέσω αριθμητικών ολισθήσεων).

Κατά τη φάση βελτίωσης, η ανακατασκευασμένη μήτρα ΥΕ η οποία προκύπτει από την πρώτη φάση φιλτράρεται και μετασχηματίζεται κατ' επανάληψη μεταξύ των χώρων ΧΕ και ΥΕ. Ο αριθμός των επαναλήψεων εξαρτάται ισχυρά από το αποτέλεσμα της φάσης δόμησης. Η φάση βελτίωσης βασίζεται σε μία αποτελεσματική τροποποίηση του πλαισίου ΕΟΠ με τη βοήθεια φίλτρων παρεμβολής χαμηλής πολυπλοκότητας. Επιπρόσθετα, η πληροφορία για την κατάσταση των ακμών η οποία προκύπτει κατά τη φάση δόμησης, προάγεται κατά τη φάση βελτίωσης μέσω της προτεινόμενης από εμάς *CxHD* τεχνικής. Η τελευταία αποτιμά την κατάσταση των ακμών σε μία τοπική περιοχή εικονοστοιχείων (αντί της στατιστικής αποτίμησης της πρώτης φάσης) προκειμένου να προσαρμοστεί στις ασυνέχειες φωτεινότητας της μετασχηματισμένης εικόνας. Συνολικά, η αλγοριθμική διαδικασία αποτελείται από 4 βήματα: *i*) υπολογισμός κατωφλίου, *ii*) χαρακτηρισμός εικονοστοιχείων, *iii*) παρεμβολή τα οποία συνιστούν την πρώτη φάση του αλγορίθμου *iv*) επαναληπτική προσαρμοστική βελτίωση σφάλματος η οποία αντιπροσωπεύει τη δεύτερη φάση του αλγορίθμου.

Έστω \mathcal{Y}^{LR} και \mathcal{Y}^{HR} ο συμβολισμός των επιπέδων φωτεινότητας ΧΕ και ΥΕ, αντίστοιχα. Για εικόνα εισόδου $W \times H$ εικονοστοιχείων, έστω $\mathcal{Y}_{i,j}^{LR}$ το ευρισκόμενο στη θέση (i, j) εικονοστοιχείο της μήτρας \mathcal{Y}^{LR} , όπου $i, j \in \mathbb{N}, i < H, j < W$. Ομοίως, με $\mathcal{Y}_{i,j}^{HR}$ συμβολίζουμε το ευρισκόμενο στη θέση (i, j) εικονοστοιχείο της μήτρας \mathcal{Y}^{HR} , όπου $i, j \in \mathbb{N}, i < f \cdot H, j < f \cdot W$, και f ο παράγοντας υπερδειγματοληψίας από τη μήτρα \mathcal{Y}^{LR} στη μήτρα \mathcal{Y}^{HR} . Για $f=2$, υποθέτουμε ότι τα εικονοστοιχεία της μήτρας \mathcal{Y}^{HR} τα οποία θα βρίσκονται σε θέσεις άρτιων συντεταγμένων αντιγράφονται από τη μήτρα \mathcal{Y}^{LR} , δηλαδή,

$$\forall i, j \in \mathbb{N}, \quad \mathcal{Y}_{2i,2j}^{HR} = \mathcal{Y}_{i,j}^{LR} \quad (5.1)$$

Τα υπόλοιπα εικονοστοιχεία της μήτρας \mathcal{Y}^{HR} ταξινομούνται σε τρία σύνολα αναλόγως της θέσης τους σε σχέση με τα "αρχικά" εικονοστοιχεία $\mathcal{Y}_{2i,2j}^{HR}$. Πιο συγκεκριμένα ακολουθούμε το διαχωρισμό των προηγούμενων κεφαλαίων, διακρίνοντας σε "ημι-οριζόντια" (ΗΟ, Half-Horizontal - HH), "ημι-κατακόρυφα" (ΗΚ, Half-Vertical - HV), και "ημι-διαγώνια" (ΗΔ, Half-Diagonal - HD) σύνολα εικο-

νοστοιχείων, οι θέσεις των οποίων ορίζονται ως:

$$\begin{aligned}\mathcal{Y}_{HH} &= \{\mathcal{Y}_{2i,2j+1}^{HR} \mid i, j \in \mathbb{N}\} \\ \mathcal{Y}_{HV} &= \{\mathcal{Y}_{2i+1,2j}^{HR} \mid i, j \in \mathbb{N}\} \\ \mathcal{Y}_{HD} &= \{\mathcal{Y}_{2i+1,2j+1}^{HR} \mid i, j \in \mathbb{N}\}\end{aligned}\quad (5.2)$$

Με βάση την ανωτέρω ονοματολογία/συμβολισμό, οι εικόνες 5.3 και 5.4 περιγράφουν τις δύο φάσεις του αλγορίθμου *L-SEABI*, συμπεριλαμβανομένων των 4 βημάτων τα οποία εξηγούμε παρακάτω.

5.2.1.1 Υπολογισμός κατώφλιου

Κατά την εκτέλεση του αλγορίθμου, χρησιμοποιούμε κατωφλίωση προκειμένου να αποφασίσουμε εάν κάποιο εικονοστοιχείο της μήτρας ΥΕ θα προκύψει από περιοχή η οποία περιλαμβάνει ακμή. Η κατωφλίωση βασίζεται στη συνολική μεταβλητότητα (Total Variation - TV) των τιμών κλίσης της εικόνας ΧΕ, και δίδεται από την Εξίσωση 5.3:

$$\mathcal{TV}^{LR} = \sum_{i,j=0}^{H,W} \mathcal{G}_x(i,j)^2 + \mathcal{G}_y(i,j)^2 \quad (5.3)$$

όπου με $\mathcal{G}_x(i,j) = \mathcal{Y}_{i,j}^{LR} - \mathcal{Y}_{i,j-1}^{LR}$, $\mathcal{G}_y(i,j) = \mathcal{Y}_{i,j}^{LR} - \mathcal{Y}_{i-1,j}^{LR}$ συμβολίζουμε τις τιμές της κλίσης στις οριζόντιες και κατακόρυφες διευθύνσεις αντιστοίχως. Το κατώφλι για κάθε εικόνα υπολογίζεται κατόπιν μέσω της Εξίσωσης 5.4, από την οποία αποθηκεύουμε την αμέραια τιμή του T :

$$T = \sqrt{\frac{\mathcal{TV}^{LR}}{2WH}} \quad (5.4)$$

5.2.1.2 Χαρακτηρισμός εικονοστοιχείων

Κατά το δεύτερο βήμα του αλγορίθμου διακρίνουμε μεταξύ εικονοστοιχείων ευρισκομένων σε περιοχές ακμών και εικονοστοιχείων τα οποία ανήκουν σε περιοχές ομοιογενούς φωτεινότητας. Η διάκριση βασίζεται α) στο υπολογισμένο κατά το προηγούμενο βήμα κατώφλι T και β) στην απόλυτη διαφορά φωτεινότητας μεταξύ δύο εικονοστοιχείων ΧΕ σε θέσεις πλησίον του προς υπολογισμό εικονοστοιχείου ΥΕ. Εάν η υπολογιζόμενη διαφορά φωτεινότητας υπερβαίνει - ή ισούται με - το κατώφλι εικόνας T , τότε θεωρούμε ότι το εικονοστοιχείο ανήκει σε ακμή. Αμέσως μετά θα περιγράψουμε το μηχανισμό ανίχνευσης ο οποίος βασίζεται στον αλγόριθμο ΤΕΔ [82], τον οποίο και έχουμε τροποποιήσει ώστε να ανιχνεύουμε ακμές με μεγαλύτερη ακρίβεια.

Στην περίπτωση εικονοστοιχείων σε ΗΟ θέσεις, συγκρίνουμε την τιμή του κατώφλιου με τη διαφορά $d_{HH} = |\mathcal{Y}_{i,j}^{LR} - \mathcal{Y}_{i,j+1}^{LR}|$ και στην περίπτωση ΗΚ θέσεων, με τη διαφορά $d_{HV} = |\mathcal{Y}_{i,j}^{LR} - \mathcal{Y}_{i+1,j}^{LR}|$. Για τις ΗΔ θέσεις χρησιμοποιούμε δύο τιμές

διαφορών, τη $d_{HD}^{SE} = |\mathcal{Y}_{i,j}^{LR} - \mathcal{Y}_{i+1,j+1}^{LR}|$ και τη $d_{HD}^{NE} = |\mathcal{Y}_{i,j+1}^{LR} - \mathcal{Y}_{i+1,j}^{LR}|$. Θεωρούμε ότι η εξεταζόμενη ΗΔ θέση περιέχει ακμή όταν και οι δύο διαφορές d_{HD}^{SE} και d_{HD}^{NE} είναι μεγαλύτερες είτε ίσες με το T . Στην περίπτωση που ισχύει $d_{HD}^{SE} < T$, τότε η περιοχή η οποία ορίζεται από τα άμεσα γειτονικά της ΗΔ θέσης εικονοστοιχεία $\mathcal{Y}_{i,j}^{LR}$, $\mathcal{Y}_{i+1,j+1}^{LR}$ θεωρείται ομοιογενής. Ομοίως, η σχέση $d_{HD}^{NE} < T$, συνεπάγεται ότι τα εικονοστοιχεία $\mathcal{Y}_{i,j+1}^{LR}$, $\mathcal{Y}_{i+1,j}^{LR}$ περικλείουν την ομοιογενή περιοχή. Τέλος, κατά την ειδική περίπτωση όπου $d_{HD}^{SE}, d_{HD}^{NE} < T$, η μικρότερη διαφορά μεταξύ των d_{HD}^{SE} και d_{HD}^{NE} προσδιορίζει και τα εικονοστοιχεία της ομοιογενούς περιοχής.

5.2.1.3 Παρεμβολή

Κατά το τρίτο βήμα του αλγορίθμου υπολογίζουμε την προκύπτουσα τιμή φωτεινότητας επιλέγοντας μεταξύ Διγραμμικής και Δικυβικής παρεμβολής, ανάλογα με το αποτέλεσμα του προηγούμενου βήματος. Χρησιμοποιούμε Δικυβική παρεμβολή με πολυωνυμική παράμετρο $\alpha = -0.75$ [46] για εικονοστοιχεία εντός ακμής. Ειδικότερα, για περιοχές ομοιογενούς φωτεινότητας χρησιμοποιούμε Διγραμμική παρεμβολή. Σε κάθε περίπτωση, η διαδικασία φιλτραρίσματος περιλαμβάνει αποθήκευση και υπολογισμούς ακεραίων αριθμών.

5.2.1.4 Προσαρμοστική οπισθοπροβολή

Στοχεύοντας σε ένα αποδοτικό αλλά απλοποιημένο σχήμα ΕΟΠ με μία μόνο εικόνα/παρατήρηση ΧΕ, στο κεφάλαιο 4 αναπτύξαμε την τεχνική προσαρμοστικής οπισθοπροβολής του σφάλματος επάνω στην εικόνα [30]. Στο τρέχον κεφάλαιο προάγουμε την ανωτέρω προσέγγιση προτείνοντας τα βήματα 5.2.1.1-5.2.1.3 για την αντικατάσταση του συμβατικού αλγορίθμου ανακατασκευής \mathcal{K} με σκοπό τη βελτίωση της ανίχνευσης τυχόν ασυνεχειών στην εικόνα. Εκτός αυτού, ο αλγόριθμος L -SEABI ενσωματώνει και βελτιώνει τις τεχνικές χαμηλής πολυπλοκότητας οι οποίες προτάθηκαν στο Κεφάλαιο 3 [29]. Υπενθυμίζουμε ότι στο κεφάλαιο 3 εισαγάγαμε τις $CxScale$ και $CrossHD$, ως προσαρμοστικές τεχνικές χαμηλής πολυπλοκότητας. Η $CxScale$ συνδυάζει Δικυβική και Διγραμμική παρεμβολή προκειμένου να υπολογίσει ένα δείγμα στο δισδιάστατο πλέγμα εικονοστοιχείων \mathcal{Y}^{HR} με διαστάσεις $2W \times 2H$. Αρχικά ανιχνεύονται ακμές με οριζόντια και κατακόρυφη διεύθυνση εντός μιας περιοχής 8 εικονοστοιχείων. Εν συνεχεία εξετάζεται ενδεχόμενη βελτίωση της αρχικής διεύθυνσης και τελικώς εφαρμόζεται Δικυβική ή Διγραμμική παρεμβολή ανάλογα με το προηγούμενο αποτέλεσμα. Η τεχνική $CrossHD$ εστιάζει στην απλοποίηση του υπολογισμού της φωτεινότητας ενός εικονοστοιχείου σε ΗΔ θέση: βασίζεται και αυτή στη φωτεινότητα 8 γειτονικών δειγμάτων, προκειμένου να εφαρμόσει με συνθηκοταγή τρόπο ένα Διγραμμικό πυρήνα σε δύο εικονοστοιχεία γειτονικά του \mathcal{Y}_{HD} . Εκτός αυτού, προτείναμε τη συγχώνευση των δύο τεχνικών αυτών (δηλαδή τον υπολογισμό των δειγμάτων \mathcal{Y}_{HH} , \mathcal{Y}_{HV} μέσω της $CxScale$ και των δειγμάτων \mathcal{Y}_{HD} μέσω της $CrossHD$) στην τεχνική $CxHD$ [30] ως έναν ακμοστρεφή και αποδοτικό αλγόριθμο υπερ-δειγματοληψίας της μήτρας σφάλματος.

- L-SEABI Φάση Δόμησης:** βήματα 1–3
- Είσοδος:** \mathcal{Y}^{LR} // εικόνα χαμηλής ευκρίνειας $W \times H$
εικονοστοιχείων
- Έξοδος:** \mathcal{Y}^{HR} // αρχική $2W \times 2H$ ανακατασκευή
- 1: για κάθε $\mathcal{Y}_{i,j}^{LR}$ εκτέλεσε
 - 2: υπολογισμός ανάδελτα $\mathcal{G}_x(i, j)$, $\mathcal{G}_y(i, j)$
 - 3: χρήση ανάδελτα στην άθροιση \mathcal{TV}^{LR} (Εξ. 5.3)
 - 4: ολοκλήρωση βρόχου
 - 5: υπολογισμός κατωφλίου T (Εξ. 5.4)
 - 6: για κάθε $\mathcal{Y}_{i,j}^{LR}$ εκτέλεσε
 - 7: αντιγραφή $\mathcal{Y}_{i,j}^{LR} \rightarrow \mathcal{Y}_{2i,2j}^{HR}$
 - 8: υπολογισμός διαφορών $d_{HH}, d_{HV}, d_{HD}^{SE}, d_{HD}^{NE}$
 - 9: εάν $d_{HH} \geq T$,
 $\mathcal{Y}_{HH} \leftarrow$ δικυβική παρεμβολή
 - 10: διαφορετικά,
 $\mathcal{Y}_{HH} \leftarrow$ διγραμμική παρεμβολή
 - 11: εάν $d_{HV} \geq T$,
 $\mathcal{Y}_{HV} \leftarrow$ δικυβική παρεμβολή
 - 12: διαφορετικά,
 $\mathcal{Y}_{HV} \leftarrow$ διγραμμική παρεμβολή
 - 13: εάν $(d_{HD}^{SE}, d_{HD}^{NE} < T$ και $d_{HD}^{NE} < d_{HD}^{SE})$ ή $(d_{HD}^{NE} < T)$
 $\mathcal{Y}_{HD} \leftarrow$ διγραμμική παρεμβολή με χρήση $\mathcal{Y}_{i,j+1}^{LR}$, $\mathcal{Y}_{i+1,j}^{LR}$
 - 14: εάν $(d_{HD}^{SE}, d_{HD}^{NE} < T$ και $d_{HD}^{SE} < d_{HD}^{NE})$ ή $(d_{HD}^{SE} < T)$
 $\mathcal{Y}_{HD} \leftarrow$ διγραμμική παρεμβολή με χρήση $\mathcal{Y}_{i,j}^{LR}$, $\mathcal{Y}_{i+1,j+1}^{LR}$
 - 15: διαφορετικά,
 $\mathcal{Y}_{HD} \leftarrow$ δικυβική παρεμβολή
 - 16: ολοκλήρωση βρόχου
επέστρεψε ανακατασκευή \mathcal{Y}^{HR}

Εικόνα 5.3: Τα βήματα 1-3 του αλγορίθμου Υπερ-Ανάλυσης L-SEABI.

Στο παρόν κεφάλαιο, χρησιμοποιούμε μία τροποποιημένη έκδοση της τεχνικής *CxHD* για τη διαδικασία υπερ-δειγματοληψίας του σφάλματος. Αναλυτικότερα, τροποποιούμε την τεχνική *CxScale* ούτως ώστε να α) υπολογίζει το μέσο όρο των τιμών φωτεινότητας για τους συνθηκοταγείς ελέγχους κλίσης και β) επιλέγονται προσαρμοστικά οι τιμές των σταθερών της Δικυβικής παρεμβολής. Η επιλογή των σταθερών εξαρτάται από τον προσανατολισμό της ανιχνευθείσας ακμής: τα εικονοστοιχεία σε ακμές με *NE* διεύθυνση, υπολογίζονται από Δικυβική παρεμβολή με πολυωνυμική παράμετρο $a = -1$, τα εικονοστοιχεία σε ακμές με *SE* διεύθυνση, με $a = -0.75$ και τα εικονοστοιχεία σε οριζόντες ή κατακόρυφες ακμές με $a = -0.5$. Κατά την οπισθοπροβολή, χρησιμοποιούμε τη μήτρα G διαστάσεων 5×5 εντεινώντας την προσοχή μας στη χρήση ακεραίων συντελεστών [30].

L-SEABI Φάση Βελτίωσης: προσαρμοστική ΕΟΠ
Είσοδος: \mathcal{Y}^{HR} // Αποτέλεσμα Φάσης Ανακατασκευής
Έξοδος: ${}^n\mathcal{Y}^{HR}$ // εικόνα υψηλής ευκρίνειας
Προκαθορισμός: σταθερές $G, H \in \mathbb{R}^{5 \times 5}$ και $n, f \in \mathbb{N}$
Προκαθορισμός: $error = \infty$, αλγόριθμος $\mathcal{L} \leftarrow \text{CxHD}$

- 1: για $p = 0$ έως $n - 1$ εκτέλεσε
- 2: συνέλιξη ${}^p\mathcal{Y}^{HR}$ με H και αποτέλεσμα $H * {}^p\mathcal{Y}^{HR}$
- 3: αποδεκίαση ($H * {}^p\mathcal{Y}^{HR}$) και αποτέλεσμα ${}^p\mathcal{Y}^{LR}$:
 ${}^p\mathcal{Y}^{LR} \leftarrow \downarrow_f (H * {}^p\mathcal{Y}^{HR})$
- 4: αφαίρεση ${}^p\mathcal{Y}^{LR}$ from \mathcal{Y}^{LR} και αποτέλεσμα ${}^p\mathcal{E}^{LR}$:
 ${}^p\mathcal{E}^{LR} \leftarrow \mathcal{Y}^{LR} - {}^p\mathcal{Y}^{LR}$
- 5: υπολογισμός σφάλματος e^i μέσω l_1 norm :
 $e^i \leftarrow \|{}^p\mathcal{E}^{LR}\|_1$
- 6: **εάν** $e^i < error$ **τότε**
- 7: υπερ-δειγματοληψία ${}^p\mathcal{E}^{LR}$ κατά f χρησιμοποιώντας τον προσαρμοστικό σε ακμές αλγόριθμο \mathcal{L} και αποτέλεσμα ${}^p\mathcal{E}^{HR}$:
 ${}^p\mathcal{E}^{HR} \leftarrow \uparrow_f^{\mathcal{L}} ({}^p\mathcal{E}^{LR})$
- 8: συνέλιξη ${}^p\mathcal{E}^{HR}$ με G και αποτέλεσμα $G * {}^p\mathcal{E}^{HR}$
- 9: βελτίωση ${}^p\mathcal{Y}^{HR}$ με $G * {}^p\mathcal{E}^{HR}$ και αποτέλεσμα ${}^{p+1}\mathcal{Y}^{HR}$:
 ${}^{p+1}\mathcal{Y}^{HR} \leftarrow {}^p\mathcal{Y}^{HR} + G * {}^p\mathcal{E}^{HR}$
- 10: **διαφορετικά επέστρεψε** ${}^p\mathcal{Y}^{HR}$ // πρόωρος τερματισμός
- 11: ανανέωση: $error \leftarrow e^i$
- 12: **ολοκλήρωση βρόχου**
- 13: **επέστρεψε** ${}^n\mathcal{Y}^{HR}$

Εικόνα 5.4: Το 4ο βήμα του αλγορίθμου L-SEABI.

5.2.2 Αξιολόγηση αλγορίθμων ΥΑ και του L-SEABI

Στις ακόλουθες παραγράφους αξιολογούμε την απόδοση και το κόστος των ανωτέρω αλγορίθμων ΥΑ όταν εφαρμόζονται σε ακίνητες εικόνες, απαλλαγμένες από οποιοδήποτε είδος θορύβου μπορεί να εισαχθεί από τυχόν ενδιάμεσες διεργασίες (π.χ. απωλεστική συμπίεση). Για λόγους συνέπειας, χρησιμοποιούμε το ίδιο σύνολο εικόνων με αυτό της ενότητας 5.1. Αναλυτικά, το σύνολο των δοκιμών μας περιλαμβάνει 26 ακίνητες εικόνες με τα εξής χαρακτηριστικά {Εικονογραφίες}/{Ανάλυσης}: {*Carphone*}/{176×144}, {*Bridge, Bus, Coastguard, Foreman, News*}/{352×288}, {*Blue Sky, Pedestrian, Riverbed, Rush Hour*}/{720×576, 1280×720, 1920×1088}, {*Shields*}/{1280×720} και τις {*Cactus, Coastguard, Foreman, Mobile, News, Park, Suzie*}/{3840×2160}. Κάθε δοκιμή περιλαμβάνει την αποδεκία της εισόδου κατά παράγοντα 2 σε κάθε διεύθυνση, την υπερδειγματοληψία της εικόνας ΧΕ κατά 2 σε κάθε διεύθυνση μέσω των υπό εξέταση αλγορίθμων¹ και τελικώς, την αποτίμηση της ποιότητας των υπερδειγματοληπτημένων εικόνων μέσω αντικειμενικών και υποκειμενικών συγκρίσεών τους με τις εικόνες αναφοράς.

Ακολουθούμε μία διαδικασία διερεύνησης-αποτίμησης χωρισμένη σε τρία στάδια, τα αποτελέσματα της οποίας παρατίθενται χωριστά ακολούθως. Στο πρώτο στάδιο, διερευνούμε την ποιότητα συνηθισμένων μεθόδων αποδεκίας με σκοπό να προσδιορίσουμε την πλέον κατάλληλη για την ενίσχυση της ποιότητας του αλγορίθμου L-SEABI. Μεταβάλλουμε λοιπόν τις παραμέτρους αποδεκίας μετρώντας την ποιότητα της προκύπτουσας εικόνας αφού εφαρμοστεί είτε ο L-SEABI είτε απλά φίλτρα παρεμβολής. Εν συνεχεία, στο δεύτερο στάδιο χρησιμοποιούμε την προσδιορισμένη μέθοδο αποδεκίας προκειμένου να μετρήσουμε και να συγκρίνουμε την ποιότητα του L-SEABI με εκείνη όλων των υπολοίπων αλγορίθμων υπερδειγματοληψίας. Τέλος, στο τρίτο στάδιο αποτιμούμε την πολυπλοκότητα των αλγορίθμων ΥΑ.

Αλγόριθμος αποδεκίας

Στο στάδιο αυτό της διερεύνησης φιλτράρουμε την εικόνα με ιστροπικές γκαουσιανές διαχωρίσιμες ΣΔΣ. Έπειτα διερευνούμε και προσδιορίζουμε διαδοχικά α) την ακτίνα β) τη συνθήκη απόκλιση και γ) τη γραμμικότητα του φίλτρου αποδεκίας. Η τελική επιλογή βασίζεται στην ποιότητα της ανακατασκευασμένης εικόνας ΥΕ και το χρόνο ολοκλήρωσης της διαδικασίας φιλτραρίσματος.

Αρχικά θέλουμε να καθορίσουμε το μέγεθος του πυρήνα συνέλιξης (αριθμός δειγμάτων φίλτρου). Για κάθε μέγεθος πυρήνα, εξετάζουμε προσωρινά ορισμένες

¹Κατά τις δοκιμές μας χρησιμοποιούμε τον MATLAB κώδικα των συγγραφέων, όπως αυτός διατέθηκε παρακάτω:

[98]: www.ifp.illinois.edu/~jyang29/ScSR.htm

[23]: www4.comp.polyu.edu.hk/~cslzhang/ASDS_AReg.htm

[90] [89]: decsai.ugr.es/pi/superresolution/software.html.

επίσης, ο αλγόριθμος L-SEABI αναπτύχθηκε και αξιολογήθηκε σε περιβάλλον MATLAB για τις ανάγκες της ενότητας 5.2.

τιμές απόκλισης σε αδρομερή βήματα μεταβάλλοντας την τιμή του σ_d από 0.2 έως 1.6 (η λεπτομερής ρύθμιση του σ_d διενεργείται μετά τον καθορισμό των διαστάσεων του πυρήνα συνέλιξης). Αναφορικά με την πολυπλοκότητα, η συνέλιξη με δισδιάστατους διαχωρίσιμους πυρήνες μεγέθους $W \times H$ μειώνει τον αριθμό των πράξεων ανά δείγμα σε $W + H$. Για παράδειγμα, η συνέλιξη μίας εικόνας μεγέθους 1080p με ένα φίλτρο διαστάσεων 11×11 απαιτεί περίπου 3.1x περισσότερο χρόνο σε σχέση με ένα φίλτρο 3×3 . Από τα αποτελέσματα της διερεύνησής μας προκύπτει το γενικό συμπέρασμα ότι η συνέλιξη με γκαουσιανή μήτρα 5×5 αποδίδει καλύτερα από μια διαστάσεων 3×3 (κατά 0.3 dB όταν ακολουθεί Δικυβική παρεμβολή, και έως 0.4 dB όταν ακολουθεί ο *L-SEABI*) απαιτώντας περίπου 1.4x περισσότερο χρόνο υπολογισμού. Ωστόσο, η χρήση περισσότερων δειγμάτων στο φίλτρο μπορεί ακόμη και να μειώσει τελικά την ποιότητα (π.χ., κατά 1 dB για φίλτρο διαστάσεων 11×11). Συμπερασματικά λοιπόν, οδηγούμαστε στην επιλογή αποδεκάτισης με φίλτρο 5×5 δειγμάτων.

Κατόπιν, διερευνούμε τις υπόλοιπες δύο παραμέτρους: τη γραμμικότητα του φίλτρου αποδεκάτισης και την τιμή της συνήθους απόκλισης. Οι συγγραφείς στο [87] έδειξαν ότι από τη μία πλευρά, ο συνδυασμός φιλτραρίσματος πεδίου και φιλτραρίσματος εύρους διατηρεί τις ακμές κατά το μετασχηματισμό εικόνων - εν αντιθέσει με το συμβατικό φιλτράρισμα πεδίου. Από την άλλη πλευρά, οι μη βελτιστοποιημένες υλοποιήσεις ενός μη-γραμμικού αμφίπλευρου φίλτρου για έναν πυρήνα διαστάσεων $W \times H$ έχουν πολυπλοκότητα $O(W \cdot H)$ για κάθε εικονοστοιχείο εισόδου. Στη δική μας περίπτωση, προκειμένου να εκτιμήσουμε τη δυνατότητα εφαρμογής αμφίπλευρου φίλτρου κατά την αποδεκάτιση, υλοποιήσαμε μία συνάρτηση αμφίπλευρου φιλτραρίσματος με φίλτρα 5×5 δειγμάτων και γκαουσιανές ΣΔΣ. Η εικόνα ΥΕ υπέστη επεξεργασία μέσω της συνάρτησης αυτής και στη συνέχεια αποδεκατίστηκε κατά 2 σε κάθε διεύθυνση. Για τον καθορισμό των βέλτιστων παραμέτρων αποδεκάτισης, θέσαμε τη συνήθη απόκλιση (σ_d) της συνάρτησης εγγύτητας στην τιμή 0.6 και μεταβάλλαμε την τιμή της συνήθους απόκλισης της συνάρτησης ομοιότητας από 0.1 σε 1.0 με βήμα 0.1 και από 1.0 σε 10.0 με βήμα 1.0. Η προκύπτουσα εικόνα υπερ-δειγματοληπτήθηκε και συγκρίθηκε με την εικόνα αναφοράς. Τα αποτελέσματα παρουσιάζουν μία γενικευμένη μέση μείωση της ποιότητας, η οποία κυμαίνεται από 0.118 ($\sigma_r = 10$) έως 0.353dB ($\sigma_r = 0.1$) σε σχέση με το συμβατικό φιλτράρισμα.

Κατά συνέπεια, έχοντας αποφασίσει να αποφύγουμε το αμφίπλευρο φιλτράρισμα (θέτοντας $\sigma_r = \infty$), βελτιώνουμε την απόκλιση σ_d μεταβάλλοντας τις τιμές της με βήμα 0.1. Από τη διαδικασία αυτή προέκυψε πως η τιμή $\sigma_d = 0.6$ παρέχει τη μέγιστη μέση αντικειμενική ποιότητα (έως 0.4 dB εν συγκρίσει με την τιμή $\sigma_d = 1.0$ για τις απλές μεθόδους, και άνω του 1 dB με τη χρήση του *L-SEABI*).

Αξιολόγηση ποιότητας αλγορίθμων υπερδειγματοληψίας

Έχοντας καθορίσει τη βέλτιστη μέθοδο αποδεκάτισης πριν την εκτέλεση του *L-SEABI* (με γκαουσιανή ΣΔΣ διαστάσεων 5×5 και $\sigma_d = 0.6$, την οποία υιοθετούμε και εντός του αλγορίθμου *L-SEABI* κατά την αποδεκάτιση της φάσης βελτιώ-

σης), προχωρούμε στην αποτίμηση της απόδοσής του ως αλγορίθμου ΥΑ. Κατά τις δοκιμές μας, κατηγοριοποιούμε τις μεθόδους ΥΑ με βάση τη σχετική τους πολυπλοκότητα, την ποιότητα των αποτελεσμάτων και το υποκείμενο πλαίσιο λειτουργίας τους, ως εξής: απλές μέθοδοι (π.χ. Δικυβική), παραπλήσιες με τον *L-SEABI* (π.χ. *EOII*-παγείς), λεξικοπαγείς (π.χ., Yang κ.ά. [98]) και πρότερης γνώσης (π.χ., Norm l1 [90]). Όλοι οι αλγόριθμοι δοκιμάστηκαν χρησιμοποιώντας τις προταθείσες από τους αντιστοίχους συγγραφείς παραμέτρους. Αρχικά, παρουσιάζουμε τα αποτελέσματα των αντικειμενικών μετρικών ποιότητας. Προκειμένου να διευρύνουμε την προοπτική μας και να καταλήξουμε σε μια όσο το δυνατόν δικαιότερη σύγκριση, υιοθετούμε δύο επιπλέον μετρικές ποιότητας εκτός του PSNR. Τη μετρική μέσης δομικής ομοιότητας (Mean Structural SiMilarity - MSSIM [93]) και τη μετρική BRISQUE (Blind Referenceless Image Spatial Quality Evaluator [57]). Η τελευταία αποτιμά τη “φυσικότητα” των υπερ-δειγματοληπτημένων εικόνων βασιζόμενη σε εμπειρικές κατανομές και στατιστικά μοντέλα φυσικών τοπίων. Επίσης, σε αντίθεση με το PSNR και την MSSIM, δεν απαιτεί αρχική εικόνα αναφοράς.

Στον πίνακα 5.1 παρατίθενται οι μέσες τιμές των μετρικών PSNR, MSSIM, και BRISQUE ανά μέγεθος εικόνας και αλγόριθμο. Η επίτευξη υψηλής ποιότητας σηματοδοτείται από υψηλές τιμές PSNR, MSSIM και χαμηλές τιμές BRISQUE. Τα αποτελέσματα δείχνουν ότι η απλή Δικυβική παρεμβολή παράγει τις χαμηλότερες τιμές PSNR και MSSIM σε σχέση με όλες τις υπόλοιπες κατηγορίες μεθόδων. Από την αποτίμηση παραπλήσιων με την προτεινόμενη μεθόδων, συμπεραίνουμε ότι τόσο το *L-SEABI* όσο και η *MTEOII (NLBP)* [22] μέθοδος βελτιώνουν το συμβατικό σχήμα *EOII* [40] όπως καταγράφεται από όλες τις μετρικές. Συν τοις άλλοις, ο αλγόριθμος *L-SEABI* υπερσιχύει της μεθόδου *NLIBP* [22] όπως επίσης καταγράφεται από όλες τις μετρικές, και συνεπώς, αποτελεί την επικρατέστερη από τις παραπλήσιες μεθόδους.

Συγκρινόμενος με τις λεξικοπαγείς μεθόδους, ο προτεινόμενος αλγόριθμος επιτυγχάνει συγκρίσιμα αποτελέσματα ποιότητας με την αραιή αναπαράσταση [98], σε όλες τις εξεταζόμενες μετρικές. Οι υψηλότερες τιμές PSNR και MSSIM καταγράφονται από τον αλγόριθμο Αγκυροβολημένης Παλινδρόμησης Γειτονίας (Anchored Neighborhood Regression, ANR) και σε εικόνες μεγέθους 1080p από τη μέθοδο Προσαρμοστικής Κανονικοποίησης Μη-Τοπικής Ομοιότητας (Nonlocal Similarity Adaptive Regularization - ASDS-AR-NL) [23]. Τα αποτελέσματα αυτά ήταν αναμενόμενα εξαιτίας της αυξημένης στιβαρότητας (robustness) του κριτηρίου κανονικοποίησης. Να σημειώσουμε εδώ ότι ο *L-SEABI* εξακολουθεί να παράγει παραπλήσια MSSIM αποτελέσματα και με τους αλγορίθμους αυτούς. Εξετάζοντας τα αποτελέσματα μέσω της BRISQUE μετρικής, βλέπουμε ότι ο *L-SEABI* αποδίδει σχεδόν σε όλες τις περιπτώσεις καλύτερα από τις τρεις λεξικοπαγείς μεθόδους.

Οι μέθοδοι πρότερης γνώσης [90] επιτυγχάνουν την ανώτερη δυνατή ποιότητα όταν χρησιμοποιούνται περισσότερες από μία εικόνες εισόδου. Αναλυτικότερα, όπως φαίνεται από τον πίνακα 5.1, όταν γεννώνται 8 εικόνες εισόδου η διαφορά ΔPSNR της μεθόδου [90] σε σχέση με την [86] είναι σημαντική: κυμαίνεται από

Πίνακας 5.1: Αντικειμενική σύγκριση ποιότητας του προτεινόμενου αλγορίθμου *L-SEABI* έναντι της Δικυβικής παρεμβολής και υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f=2$). Όλοι οι αλγόριθμοι έχουν εκτελεστεί με τις προτεινόμενες από τους συγγραφείς τους παραμέτρους. Η χαμηλότερη τιμή της μετρικής BRISQUE αντιστοιχεί σε υψηλότερη ποιότητα εικόνας.

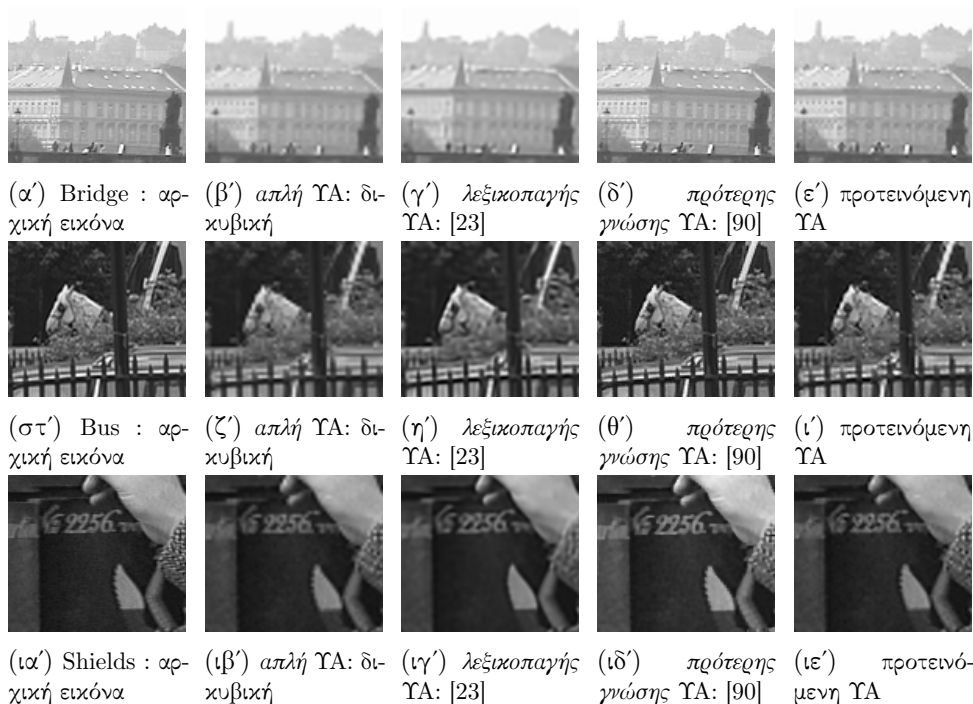
Μέθοδος: μετρική ανάλυση	Δικυβική		IBP		Προτεινόμενη		NLBP		ANR		Yang et al.		ASDS-AR-NL		Norm l_1	
	Matlab	[40]	[22]	[22]	[86]	[98]	[23]	[90]								
PSNR (dB)	QCIF	28.4333	29.5487	32.1817	32.0960	32.8185	32.7391	32.6649	37.6142							
	CIF	27.6865	29.3236	30.3245	30.6781	32.2173	31.4638	30.8917	33.9366							
	SD1	34.7955	37.4003	40.1361	39.2954	41.8011	40.7740	40.4893	-							
	720p	35.2215	38.2231	40.2407	39.4268	41.8233	40.9537	40.6771	-							
	1080p	39.0407	42.9396	44.2010	42.9759	45.6160	45.0176	48.7362	-							
	2160p	43.2794	46.5114	48.6471	46.5688	50.2350	49.0415	-	-							
μ.ό.	34.7428	37.3244	39.2885	38.5068	40.7519	39.9983	-	-	-							
MSSIM	QCIF	0.8653	0.8745	0.9114	0.9025	0.9096	0.9169	0.9040	0.9671							
	CIF	0.8253	0.8545	0.8763	0.8632	0.8797	0.8780	0.8552	0.9609							
	SD1	0.9523	0.9664	0.9702	0.9612	0.9762	0.9737	0.9687	-							
	720p	0.9407	0.9549	0.9403	0.9502	0.9817	0.9614	0.9572	-							
	1080p	0.9719	0.9798	0.9796	0.9721	0.9832	0.9811	0.9779	-							
	2160p	0.9833	0.9875	0.9880	0.9821	0.9896	0.9880	-	-							
μ.ό.	0.9231	0.9363	0.9474	0.9386	0.9533	0.9499	-	-	-							
BRISQUE	QCIF	40.2929	33.3055	39.9237	35.7802	35.8497	31.6656	40.7779	25.7769							
	CIF	40.6259	41.8366	33.8815	39.9493	41.8167	42.2242	40.1209	25.4169							
	SD1	33.6510	31.8569	27.8568	29.3260	30.5488	30.7540	33.4959	-							
	720p	46.5934	44.4594	41.6382	42.5637	41.3031	42.2648	49.3731	-							
	1080p	49.6061	48.2574	40.4554	42.3395	43.6311	45.7965	53.3932	-							
	2160p	58.7316	56.8478	56.8680	53.4940	54.4242	56.7518	-	-							
μ.τ.	44.9169	42.7608	40.1054	40.5754	41.2623	41.5761	-	-	-							

1.719 dB σε εικόνες CIF, έως 4.795 dB (για την εικόνα αναφοράς *carphone* μεγέθους 176×144). Αυτά τα αποτελέσματα θεωρούνται αναμενόμενα, εφόσον η αύξηση του αριθμού των παρατηρήσεων ΧΕ ευνοεί το κακώς τεθειμένο πρόβλημα ΥΑ. Σε αυτό το σημείο καλούμε επίσης τον αναγνώστη να παρατηρήσει ότι στην περίπτωση μίας εικόνας εισόδου (όπως έχει σχεδιαστεί ο *L-SEABI*), οι τιμές του PSNR για τη μέθοδο [90] υποβαθμίζονται σημαντικά, π.χ. 29.6788 dB σε εικόνες CIF, δηλαδή χαμηλότερα από το *L-SEABI*.

Με την εικόνα 5.5 παρέχεται επιπλέον η δυνατότητα υποκειμενικής σύγκρισης των πιο διακεκριμένων μεθόδων ΥΑ σε κάθε μία από τις εξεταζόμενες κατηγορίες, με βάση την PSNR μετρική. Παρατηρούμε ότι η απλή Δικυβική παρεμβολή (5.5β', 5.5ζ', 5.5ιβ') παρουσιάζει σημαντικά φαινόμενα θόλωσης και αναδίπλωσης εν συγκρίσει με τις υπόλοιπες μεθόδους. Όσον αφορά τις *λεξικοπαγείς* μεθόδους, η προσαρμοστική επιλογή/κανονικοποίηση του χώρου αραιής αναπαράστασης (ASDS-AR-NL) εξαλείφει τα φαινόμενα αναδίπλωσης, ωστόσο αφαιρεί και λεπτομερειακή πληροφορία, όπως φαίνεται στις εικόνες 5.5γ', 5.5η' και 5.5ιγ'. Οπτικά αυτό μεταφράζεται σε ένα σχετικά θολό αποτέλεσμα (κάτι που επιβεβαιώνεται και από τη μετρική BRISQUE). Να σημειώσουμε ωστόσο ότι αυτού του τύπου η συμπεριφορά είναι λιγότερο εμφανής με τη μέθοδο αραιής αναπαράστασης των Yang κ.ά. [98]. Η αντιπροσωπευτική των μεθόδων *πρότερης γνώσης* Norm l1 [90], διατηρεί την πλησιέστερη προσέγγιση στην ποιότητα των εικόνων αναφοράς (βλ. 5.5δ', 5.5θ', 5.5ιδ'), αποτέλεσμα το οποίο είναι σε συνέπεια με τις τιμές των μετρικών του πίνακα 5.1. Τέλος, εν συγκρίσει με τις μεθόδους μίας εικόνας, ο προτεινόμενος αλγόριθμος (εικόνες 5.5ε', 5.5ί' και 5.5ιε') διατηρεί πολύ σημαντικό τμήμα από τις λεπτομέρειες των εικόνων αναφοράς (5.5α', 5.5στ', 5.5ια').

Αξιολόγηση πολυπλοκότητας αλγορίθμων ΥΑ

Προκειμένου να αποτιμηθεί η πολυπλοκότητα των υπό εξέταση αλγορίθμων ΥΑ, μετρήσαμε το χρόνο εκτέλεσης κάθε συνάρτησης υπερ-δειγματοληψίας με δεδομένο παράγοντα $f=2$. Όλες οι δοκιμές εκτελέστηκαν σε επεξεργαστή γενικού σκοπού x86-64 αρχιτεκτονικής (8-πύρηνο, συχνότητα λειτουργίας 3.6 GHz, κεντρική μνήμη RAM 16 GB) και τα αποτελέσματα παρουσιάζονται στον πίνακα 5.2. Για τον προτεινόμενο αλγόριθμο απαιτούνται χρόνοι οι οποίοι κυμαίνονται από 75 ms σε QCIF αναλύσεις, έως 4.827 sec για εικόνες μεγέθους 2160p. Σε σύγκριση με την προτεινόμενη λύση, οι μέθοδοι αραιής αναπαράστασης απαιτούν έως και τρεις τάξεις μεγέθους περισσότερο χρόνο. Μεταξύ των *λεξικοπαγών* μεθόδων, η ANR [86] επιτυγχάνει ιδιαίτερα σημαντική μείωση του χρόνου σε σχέση με τις [98, 23], ωστόσο εξακολουθεί να χρειάζεται μία τάξη μεγέθους περισσότερο χρόνο σε σύγκριση με τον *L-SEABI*. Συν τοις άλλοις, κατά την εκτέλεση των μεθόδων αυτών υποθέτουμε ότι το λεξικό προϋπάρχει, δηλαδή ο χρόνος κατάρτισής του δε συνυπολογίζεται ως μέτρο πολυπλοκότητας (π.χ., 1768 δευτερόλεπτα για ένα λεξικό 1024 ατόμων στη μέθοδο ANR). Για να επιτύχουμε δίκαιη σύγκριση κατά την ανακατασκευή με τη μέθοδο [90] πολλαπλών εικόνων εισόδου λαμβάνουμε επιπλέον υπόψιν μας το χρόνο που απαιτείται για την εκτίμηση



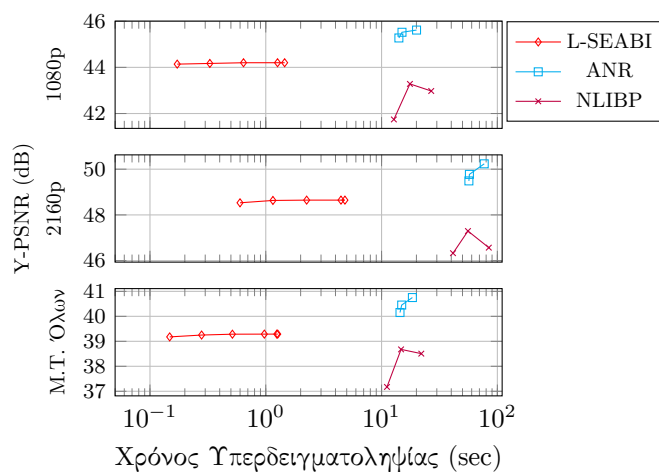
Εικόνα 5.5: Αντιπαράθεση των πιο διακεκριμένων ποιοτικά αλγορίθμων κάθε κατηγορίας για υποκειμενική σύγκριση, $f = 2$.

των παραμέτρων καταχώρισης. Από τις μετρήσεις προκύπτει ότι η ανωτέρω μέθοδος *πρότερης γνώσης* απαιτεί τουλάχιστον 4 τάξεις μεγέθους περισσότερο χρόνο από την προτεινόμενη (η ανακατασκευή της εικόνας *bridge* με 8 εικόνες εισόδου χρειάστηκε 11777.040 sec). Ακόμη και με προϋπολογισμένες τις παραμέτρους του μοντέλου αρχικής υποβάθμισης εικόνας, η βέλτιστη προσέγγιση της εικόνας αναφοράς παραμένει 10^4 φορές πιο βραδεία από τον προτεινόμενο αλγόριθμο.

Από την ξεχωριστή αξιολόγηση ποιότητας και πολυπλοκότητας (Πίνακες 5.1, 5.2) προκύπτει το συμπέρασμα ότι οι μέθοδοι *L-SEABI*, *NLIBP* και *ANR* παράγουν τη μέγιστη δυνατή ποιότητα με το μικρότερο χρόνο εκτέλεσης. Στο σημείο αυτό, επεκτείνουμε περαιτέρω την αξιολόγησή μας και μελετάμε από κοινού ποιότητα και πολυπλοκότητα, μεταβάλλοντας τις παραμέτρους των μεθόδων που διακρίθηκαν. Ειδικά για τον αλγόριθμο *L-SEABI* και τη μέθοδο *NLIBP*, μεταβάλλουμε τον αριθμό των επαναλήψεων μεταξύ της μέγιστης απαιτούμενης τιμής για σύγκλιση και μίας μόνο επανάληψης (οι συγγραφείς της μεθόδου *NLIBP* προτείνουν ένα σταθερό αριθμό 20 επαναλήψεων). Επίσης μεταβάλλουμε το μέγεθος του λεξικού της μεθόδου *ANR* μεταξύ 16 και 1024 ατόμων. Τα αποτελέσματα παρουσιάζονται στην Εικόνα 5.6 για αναλύσεις εικόνας 1080p, 2160p και το μέσο όρο όλων των αναλύσεων. Αναφορικά με τον *L-SEABI*, τα αποτελέσματα δείχνουν ότι η αποτελεσματικότητα της φάσης δόμησης και η προσαρμοστικότητα της φάσης βελτίωσης, αρκούν για την παραγωγή αποτελεσμάτων υψηλής ποιότητας με μόνο μία επανάληψη. Οι υπόλοιπες επαναλήψεις βελτιώνουν την

Πίνακας 5.2: Σύγκριση πολυπλοκότητας: Μέσος χρόνος υπερ-δειγματοληψίας (sec) μεθόδων ΥΑ ($f = 2$) και μέση τιμή ανά εικονοστοιχείο (μ.τ./ε.σ.).

Μέθοδος:	Δικτυβική Matlab	ΕΟΠ [40]	Προτεινόμενη	NLIBP [22]	ANR [86]	Yang et al. [98]	ASDS-AR-NL [23]	Norm l_1 [90]
ανάλυση				Χρόνος Υπερδειγματοληψίας (sec)				
QCIF	0.0023	0.4862	0.0733	1.0857	0.2271	41.8506	102.7923	5964.8773
CIF	0.0035	1.4679	0.1648	1.7967	0.9560	181.8234	382.0333	11950.2998
SD1	0.0076	3.9701	0.3269	5.6981	3.9512	931.4906	1446.7916	-
720p	0.0116	6.9702	0.6492	12.4345	8.5374	1917.3890	3062.7306	-
1080p	0.0219	12.1274	1.4043	26.8469	20.0519	4814.0392	6832.9518	-
2160p	0.0853	61.9599	4.4499	84.3536	77.0477	15031.6495	-	-
μ.τ./ε.σ.	0.01 ·10 ⁻⁶	7.34 ·10 ⁻⁶	0.59 ·10 ⁻⁶	11.16 ·10 ⁻⁶	9.3926 ·10 ⁻⁶	1934.61 ·10 ⁻⁶	3329.76 ·10 ⁻⁶	6652.42 ·10 ⁻⁴



Εικόνα 5.6: Από κοινού αποτίμηση ποιότητας/πολυπλοκότητας αλγορίθμων ΥΑ.

ανακατασκευή κατά 0.11 dB οδηγώντας στα αποτελέσματα του πίνακα 5.1. Η εκτέλεση μίας μόνο επανάληψης, επιτρέπει στον αλγόριθμο *L-SEABI* να συγκλίνει ταχύτερα κατά 8.57x σε σχέση με το χρόνο του πίνακα 5.2 (η επιτάχυνση κυμαίνεται από 11.21x για εικόνες CIF έως 8x για εικόνες 2160p). Εφεξής, θα αναφερόμαστε στην εκτέλεση μίας επανάληψης του προτεινόμενου αλγορίθμου ως *SIL-SEABI*. Παρόμοια αξιολόγηση της μεθόδου *NLIBP* καταδεικνύει αύξηση στην τιμή του PSNR (έως 0.72 dB σε εικόνες UHD εικόνες και 0.17 dB κατά μέσο όρο) θέτοντας ένα σταθερό αριθμό 8 επαναλήψεων. Αυτό οφείλεται σε ένα μερικό μετριασμό των προκαλούμενων από την *NLIBP* παραμορφώσεων, λόγω της αυξημένης θόλωσης στην ανακατασκευασμένη εικόνα (οι συγγραφείς χρησιμοποιούν γκαουσιανή $\Sigma\Delta\Sigma$ με απόκλιση $\sigma_a = 1.0$). Όπως αναμενόταν, η χρήση μίας επανάληψης ελαττώνει τις μέσες επιδόσεις του αλγορίθμου κατά 1.33 dB και έως 1.83 dB σε εικόνες μεγέθους SD1 σε σύγκριση με τον πίνακα 5.1. Από την πλευρά της πολυπλοκότητας, η μέθοδος *NLIBP* συγκλίνει ταχύτερα κατά 1.49x με 8 επαναλήψεις και κατά 1.98x με μία επανάληψη (αντί των 20 του Πίνακα 5.2). Η ελάττωση του μεγέθους του λεξικού υποβαθμίζει τις μέσες επιδόσεις της μεθόδου *ANR* κατά 0.3 dB (128 άτομα) έως 0.6dB (16 άτομα και συγκεκριμένα έως 0.82 dB σε εικόνες μεγέθους QCIF). Αντίθετα με τον αλγόριθμο *SIL-SEABI*, η μέση αντίστοιχη μείωση στο χρόνο εκτέλεσης είναι μάλλον αμελητέα: συγκρίνοντας με τις τιμές του πίνακα 5.2, η μέθοδος *ANR* τερματίζει ταχύτερα κατά 1.23x με λεξικό 128 ατόμων και κατά 1.27x με λεξικό 16 ατόμων. Επομένως, μετά την από κοινού αποτίμηση ποιότητας/πολυπλοκότητας, προκύπτει το συμπέρασμα πως είναι δυνατή η περαιτέρω μείωση κατά μία τάξη μεγέθους του χρόνου εκτέλεσης του αλγορίθμου *L-SEABI*, με ελάχιστες μάλιστα επιπτώσεις στην ποιότητα των εικόνων ΥΕ (Εικ. 5.6).

Συμπερασματικά, ο συνδυασμός ακεραίων υπολογισμών, μετασχηματισμών χαμηλής πολυπλοκότητας και η βελτιωμένη φάση δόμησης του *L-SEABI* οδηγούν σε σημαντικά ταχύτερη εκτέλεση εν συγκρίσει με τις υπόλοιπες μεθόδους

ΥΑ: ο προτεινόμενος αλγόριθμος είναι κατά μία τάξη μεγέθους ταχύτερος από *paraπλήσιες* μεθόδους και έως 3 τάξεις μεγέθους ταχύτερος από τις ιδιαίτερα περίπλοκες *λεξικοπαγείς* και μεθόδους *πρότερης γνώσης* (Πίνακας 5.2). Παράλληλα, η προσαρμοστικότητα του *L-SEABI* στο περιεχόμενο των εικόνων/σφαιμάτων και η διαδικασία αμοιβαίας βελτίωσης των αλληλοεξαρτώμενων διεργασιών αποδεκτικότητας/υπερ-δειγματοληψίας μας οδήγησαν σε αντικειμενικά (Πίνακας 5.1) και υποκειμενικά (Εικ. 5.5) παρατηρήσιμο κέρδος ποιότητας.

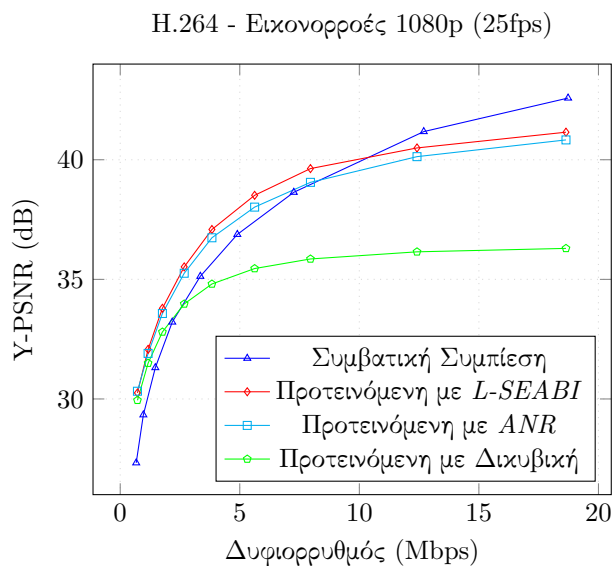
5.3 ΥΑ μέσω *L-SEABI* στο προτεινόμενο σχήμα συμπίεσης

Συνεχίζοντας την ανάλυσή μας σε αυτή την ενότητα, ύστερα από την αποτίμηση των επιδόσεων ΥΑ υιοθετούμε τον αλγόριθμο *L-SEABI* στο προτεινόμενο σχήμα συμπίεσης. Το τελικό σχήμα αξιολογείται και συγκρίνεται με το συμβατικό κωδικοποιοποιητή, το σχήμα με Δικυβική παρεμβολή, με την *paraπλήσια* μέθοδο *NLIBP* και με τη *λεξικοπαγή ANR* (η τελευταία εκτελείται με το λεξικό που παρέχουν οι συγγραφείς, χωρίς θόρυβο χβάντισης). Αξίζει να σημειωθεί εδώ ότι όπως είδαμε προηγουμένως, οι υπόλοιπες περίπλοκες μέθοδοι ΥΑ μπορούν εν δυνάμει να βελτιώσουν την ποιότητα του προτεινόμενου σχήματος. Ωστόσο, κοστίζουν υπερβολικά σε χρόνο υπολογισμού και συνεπώς, οι μέθοδοι αυτές ανήκουν εμφανώς εκτός της “εμβέλειας” χαμηλής πολυπλοκότητας της παρούσας εργασίας. Για την αποτίμηση των επιδόσεων του νέου σχήματος, αναπτύχθηκε ένα πλαίσιο λογισμικού σε γλώσσα C το οποίο ενσωματώνει τις προαναφερθείσες μεθόδους και το λογισμικό αναφοράς των προτύπων AVC και HEVC [1][5].

Για τις δοκιμές μας, ρυθμίζουμε τον κωδικοποιητή AVC ώστε να χρησιμοποιεί το Baseline προφίλ σε επίπεδο 5.0, να κωδικοποιεί μόνο το πρώτο πλαίσιο με ενδοπλαισιακή πρόβλεψη και ορίζουμε όλες τις υπόλοιπες ρυθμίσεις στις προκαθορισμένες τιμές (χωρίς έλεγχο δυφορρυθμού). Ο κωδικοποιητής HEVC χρησιμοποιεί τις ίδιες παραμέτρους ενδοπλαισιακής πρόβλεψης, ομάδα εικόνων (Group-Of-Pictures, GoP) μεγέθους 8, βάθος διαμέρισης 4 για τον κωδικοποιητή και μέγιστο μέγεθος μετασχηματισμού ίσο με 32×32 για τη μονάδα μετασχηματισμού (Transform Unit, TU). Στη μονάδα EK, θέτουμε εύρος αναζήτησης 64 εικονοστοιχείων, εύρος αμφίδρομης αναζήτησης 4 και χρήση μετασχηματισμού Hadamard κατά την EK σε υποεικονοστοιχεία. Για να παραχθεί επαρκής αριθμός σημείων στις καμπύλες δυφορρυθμού-παραμόρφωσης, μεταβάλλουμε την παράμετρο χβάντισης στο διάστημα 20 έως 49 (συμβατικός κωδικοποιητής) και 12 έως 40 (προτεινόμενο σχήμα, καθώς επιθυμούμε μικρότερο θόρυβο χβάντισης).

5.3.1 Αξιολόγηση δυφορρυθμού-παραμόρφωσης

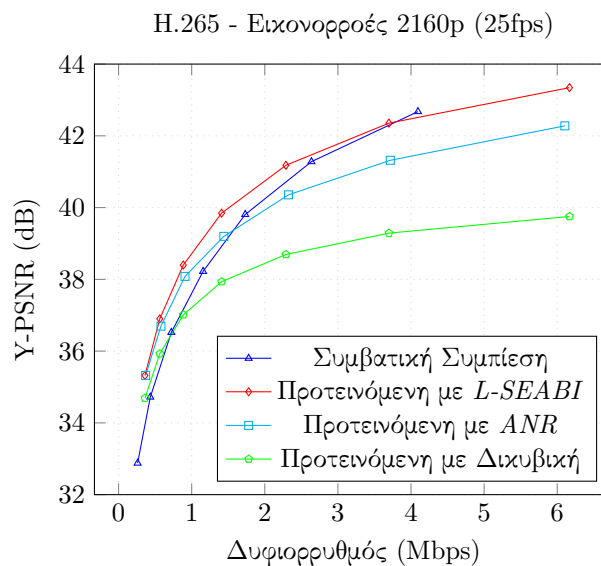
Έχοντας ως στόχο την αξιολόγηση της αποδοτικότητας των σχημάτων κωδικοποίησης, εφαρμόζουμε τα παραπάνω στο σύνολο των εικόνων της ενότητας 5.1 με αναλύσεις πλαισίων έως 1080p για το AVC και 2160p για το HEVC (με ρυθμό



Εικόνα 5.7: Βελτίωση Δυφορρυθμού-Παραμόρφωσης στο AVC μέσω του *L-SEABI*.

25 πλαίσια/sec). Καταμετρούμε το δυφορρυθμό των συμπιεσμένων εικονορροών και την PSNR μετρική των υπερ-δειγματοληπτημένων πλαισίων (uv στην Εικ. 5.1). Από τη διαδικασία προκύπτουν τα εικονιζόμενα στις 5.7 και 5.8 γραφήματα δυφορρυθμού-ποιότητας (Mbps-PSNR), για τα πρότυπα AVC και HEVC αντιστοίχως. Κάθε γράφημα περιλαμβάνει 4 καμπύλες χάριν σύγκρισης: α) με μπλε χρώμα, αυτή της συμβατικής κωδικοποίησης, β) με κόκκινο, τα αποτελέσματα του προτεινόμενου σχήματος με τον αλγόριθμο *L-SEABI*, γ) με πράσινο, τα αποτελέσματα του προτεινόμενου σχήματος με απλή Δικυβική παρεμβολή (το φίλτρο του MATLAB R2012b, υπενθυμίζουμε ως σημείο αναφοράς σε όλες τις μετρήσεις), και τέλος, δ) με κυανό, τα αποτελέσματα του προτεινόμενου σχήματος με τη μέθοδο *ANR* (για την οποία συμπεράναμε ότι αντιστοιχεί στα πιο ανταγωνιστικά αποτελέσματα χρόνου και ποιότητας σε σχέση με το *L-SEABI*).

Όπως παρουσιάσαμε και στην ενότητα 5.1, οι καμπύλες καταδεικνύουν πως υφίσταται μία τιμή μέσου “κρίσιμου” δυφορρυθμού, υπό την οποία το προτεινόμενο σχήμα επικρατεί έναντι της συμβατικής κωδικοποίησης. Εκ των αποτελεσμάτων της Εικ. 5.7 φαίνεται ότι η χρήση του *L-SEABI* μπορεί να αυξήσει σημαντικά τον κρίσιμο δυφορρυθμό από την περιοχή των 3, στα 10.4 Mbps για συμπίεση εικόνων HD. Τέτοιες τιμές δυφορρυθμών επεκτείνουν τη δυνατότητα χρήσης του προτεινόμενου σχήματος σε ποικίλλες σύγχρονες εφαρμογές, π.χ. ακόμη και σε ψηφιακή μετάδοση τηλεοπτικού σήματος (DVB-T) με QPSK ή 16-QAM σχήματα διαμόρφωσης. Ο αλγόριθμός μας βελτιώνει την ποιότητα της εικόνας του συμβατικού H.264 έως και κατά 2.67 dB για χαμηλές τιμές δυφορρυθμών (0.668 Mbps). Επίσης, παρέχει καλύτερες επιδόσεις στο προτεινόμενο σχήμα, τόσο από τη μέθοδο *NLIBP* (2.84 Mbps κρίσιμος δυφορρυθμός) όσο και από



Εικόνα 5.8: Βελτίωση Δυφορρυθμού-Παραμόρφωσης στο HEVC μέσω του *L-SEABI*.

την *ANR* (8.42 Mbps κρίσιμος δυφορρυθμός). Εκτός των άλλων, ο *L-SEABI* βελτιώνει το PSNR της Δικυβικής παρεμβολής κατά 0.31 dB στα 0.72 Mbps, και έως και κατά 5.24 dB στα 27.18 Mbps. Κατά τη μεμονωμένη εξέταση εικονορροών (ποικίλλου περιεχομένου και αναλύσεων), ο κρίσιμος δυφορρυθμός κυμάνθηκε από τα 15 Kbps στην περίπτωση του *carphone* ανάλυσης QCIF, σε 60.4 Kbps στην περίπτωση του *foreman* ανάλυσης CIF (όπου Δικυβική, *NLIBP* και *ANR* επιτυγχάνουν 51.1, 58.8 και 101 Kbps αντιστοίχως) έως την εντυπωσιακή τιμή των 31 Mbps στην περίπτωση του *Riverbed* ανάλυσης 1080p (όπου Δικυβική, *NLIBP* και *ANR* επιτυγχάνουν 8.1, 7.25 και 30 Mbps αντιστοίχως). Κατά την υπερ-δειγματοληψία μέσω του *SIL-SEABI*, δηλ., της εκτέλεσης μίας επανάληψης του *L-SEABI*, καταμετρήθηκε υποβάθμιση ποιότητας ίση με 0.1 dB στη χειρότερη περίπτωση και μέσος κρίσιμος δυφορρυθμός 9.96 Mbps. Συνεπώς, βλέπουμε ότι και κατά την εφαρμογή του *SIL-SEABI* στο σχήμα μπορούμε να επιτύχουμε καλύτερη ποιότητα από τις μεθόδους *NLIBP* και *ANR*.

Εκτός του προτύπου H.264/AVC, η μελέτη μας επιβεβαιώνει και την αποδοτικότητα του προτεινόμενου σχήματος κατά την εκτέλεση του ίδιου συνόλου δοκιμών με το σαφώς αποδοτικότερο –και περιπλοκότερο– πρότυπο HEVC (Εικ. 5.8). Εν προκειμένω, διαπιστώνουμε ότι σε εικονορροές ανάλυσης UHD, οι επιδόσεις των *NLIBP* και *ANR* εξακολουθούν να περιορίζονται από φαινόμενα παραμόρφωσης. Από την άλλη πλευρά, ο *L-SEABI* αυξάνει το μέσο κρίσιμο δυφορρυθμό στα 3.8 Mbps εν συγκρίσει με τα 0.7 Mbps της απλής Δικυβικής παρεμβολής και τα 0.59 Mbps της παραπλήσιας μεθόδου *NLIBP*. Τέλος, η χρήση της λεξικοπαγούς μεθόδου *ANR* αυξάνει το μέσο κρίσιμο δυφορρυθμό στα 1.57 Mbps. Σε συγκεκριμένες εικονορροές μεγέθους 2160p, ο *L-SEABI* αυξάνει το μέσο κρίσιμο δυφορρυθμό κατά πολύ περισσότερο από τις μέσες τιμές του γραφήμα-

Πίνακας 5.3: Επιδόσεις Σχήματος Συμπίεσης με ΥΑ.

Μέγεθος πλαisiού	NLIBP [22]	ANR [86]	L-SEABI	SIL-SEABI
	Μέσος Κρίσιμος Δυφιορρυθμός AVC (Mbps)			
SD1	0.4608	1.3899	1.3319	1.2446
720p	0.9526	3.4488	3.5070	3.4088
1080p	2.8403	8.4200	10.3938	9.9643
Μέσος Κρίσιμος Δυφιορρυθμός HEVC (Mbps)				
SD1	0.0483	0.1233	0.1069	0.1073
720p	0.1040	0.3176	0.3052	0.2430
1080p	0.2181	0.7505	0.9001	0.9069
2160p	0.5876	1.5749	3.8028	3.7658

τος 5.8: για παράδειγμα στην εικονοροή *Park*, οι καμπύλες διασταυρώνονται στην περιοχή των 13 Mbps. Το μέσο κέρδος του *L-SEABI* έναντι της Δικυβικής παρεμβολής αυξάνεται σε έως 3.5 dB στα 6.168 Mbps (και 1.04 dB στα 3.72 Mbps έναντι της *ANR*). Στην πλευρά του αποκωδικοποιητή, παρατηρούμε ότι ο *SIL-SEABI* παρέχει σχεδόν ταυτόσημα αποτελέσματα με τον *L-SEABI* εφόσον η μέση υποβάθμιση του PSNR δεν υπερβαίνει τα 0.01 dB. Εν συγκρίσει με τη συμβατική συμπίεση σε αναλύσεις UHD, το προτεινόμενο σχήμα επιτυγχάνει ένα μέσο κέρδος PSNR της τάξης των 1.4 dB στα 260 Kbps και μέγιστη τιμή 3 dB στα 99 Kbps (εικονοροή *Suzie*, όπου η *ANR* επιτυγχάνει 2.1 dB). Σε εικονοροές χαμηλότερης ανάλυσης, ο μέσος κρίσιμος δυφιορρυθμός βρίσκεται στα 0.9 Mbps για μεγέθη 1080p και 0.11 Mbps για μεγέθη SD (τα αντίστοιχα αποτελέσματα της *ANR* είναι 0.75 και 0.12 Mbps).

Συμπερασματικά, τα αποτελέσματα ποιότητας όσον αφορά την τιμή του κρίσιμου δυφιορρυθμού συνοψίζονται στον Πίνακα 5.3. Όπως ήταν αναμενόμενο, το κέρδος του προτεινόμενου σχήματος είναι υψηλότερο όταν εφαρμόζεται ο κωδικοποιητής H.264 (εφόσον το πρότυπο HEVC ελαττώνει σημαντικά το θόρυβο κβάντισης σε σύγκριση με το AVC). Τουναντίον, το H.264 χαρακτηρίζεται από σημαντικά χαμηλότερη πολυπλοκότητα σε σχέση με το HEVC, κάτι που πρέπει να ληφθεί υπόψιν σε εφαρμογές χαμηλής έως μέσης πολυπλοκότητας (δηλαδή συμβατές με την εμβέλεια της μελέτης μας).

5.3.2 Αξιολόγηση πολυπλοκότητας στο σχήμα συμπίεσης

Στην τρέχουσα ενότητα, εστιάζουμε στη συμπίεση εικονοροών μεγέθους HD και UHD προκειμένου να αποτιμήσουμε το υπολογιστικό κόστος του προτεινόμενου σχήματος. Χρησιμοποιούμε το σύνολο των εικονοροών και τον επεξεργαστή της ενότητας 5.1, με τη βοήθεια των οποίων καταμετράται ο χρόνος εκτέλεσης και η χρήση μνήμης της κάθε μονάδας του σχήματος συμπίεσης.

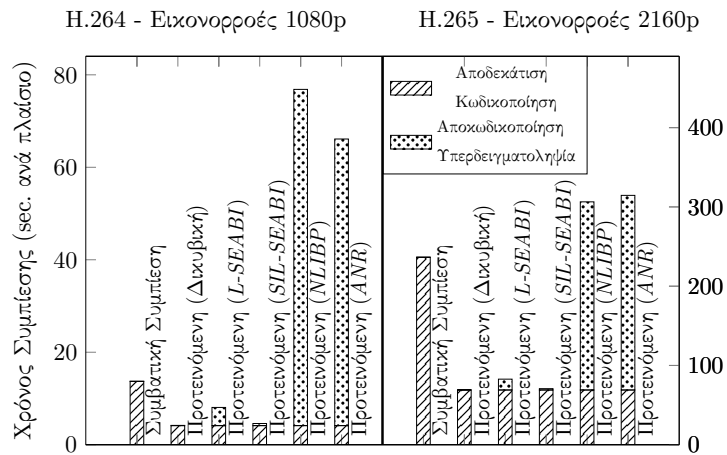
Ο κωδικοποιητής του H.264 επεξεργάζεται ένα πλαίσιο μεγέθους 1920 × 1980 εικονοστοιχείων σε 13.8 δευτερόλεπτα κατά μέσο όρο, με τυπική απόκλιση 3.6

sec ανάλογα με το περιεχόμενο της εικονορροής και την παράμετρο κβάντισης. Ο αποκωδικοποιητής χρειάζεται 0.08 sec ανά πλαίσιο. Από την άλλη πλευρά, όταν τα πλαίσια αυτά κωδικοποιούνται μέσω του προτεινόμενου σχήματος, ο χρόνος εκτέλεσης του AVC κωδικοποιητή ελαττώνεται σε περίπου 4 sec με 0.13 δευτερόλεπτα να αναλώνονται για την αποδεκίαση. Συνεπώς, επιτυγχάνεται μία επιτάχυνση κατά 3.4x στην πλευρά του κωδικοποιητή. Στην πλευρά του αποκωδικοποιητή, παρατηρείται αντίστοιχη τιμή επιτάχυνσης κατά την αποσυμπίεση των αποδεκισμένων πλαισίων, ενώ ο αλγόριθμος *L-SEABI* υπερ-δειγματοληπτεί κάθε πλαίσιο σε 3.4 sec. Άρα, το προτεινόμενο σχήμα επάγει τρόπον τινά “μεταφορά” πολυπλοκότητας από την πλευρά του κωδικοποιητή (-9.7 sec), στην πλευρά του αποκωδικοποιητή (+3.38 sec). Η μεταφορά αυτή είναι ιδιαίτερα χρήσιμη σε σχήματα “ανάστροφης πολυπλοκότητας”, ή σε παρόμοιες εφαρμογές οι οποίες απαιτούν σημαντική ελάττωση του φόρτου κωδικοποίησης χωρίς δραστικές υποβαθμίσεις στην ποιότητα της εικόνας. Συν τοις άλλοις, υποδιπλασιάζεται και ο χρόνος επεξεργασίας όλου του κωδικοαποκωδικοποιητή. Η χρήση του *SIL-SEABI* συνεπάγεται ακόμη μεγαλύτερο όφελος, μειώνοντας το χρόνο από 13.8 σε 4.6 sec, με αμελητέα ελάττωση ποιότητας σε σχέση με την υποβοηθούμενη από το *L-SEABI* αποσυμπίεση (κατά μέσο όρο μείωση PSNR 0.04 dB στα 9.96Mbps).

Θέλοντας να εξομοιώσουμε ένα υπολογιστικά απαιτητικότερο σενάριο, εξετάζουμε την απόδοση του προτεινόμενου σχήματος με το πρότυπο H.265 σε εικονορροές UHD ανάλυσης. Όπως ήταν αναμενόμενο, ο συμβατικός χρόνος κωδικοποίησης αυξάνεται κατά 17 φορές σε σχέση με τα προαναφερθέντα αποτελέσματα (δηλαδή απαιτούνται 236 δευτερόλεπτα ανά πλαίσιο). Όμοια, και η αποκωδικοποίηση με το H.265 απαιτεί 0.38 sec ανά πλαίσιο. Το προτεινόμενο σχήμα επιταχύνει την κωδικοποίηση κατά 3.4x σε πλαίσια μεγέθους 2160p (Εικ. 5.9). Όσον αφορά τον αποκωδικοποιητή, τα φίλτρα χαμηλής πολυπλοκότητας του *L-SEABI* αυξάνουν το χρόνο εκτέλεσης σχεδόν γραμμικά σε 13.6 δευτερόλεπτα/πλαίσιο. Να σημειωθεί εδώ ότι ο *L-SEABI* αντιστοιχεί στο 45% του συνολικού χρόνου με το πρότυπο H.264 και το 16% όταν εκτελείται το H.265. Τέλος, σε αυτή την περίπτωση ο *SIL-SEABI* απαιτεί 1.7 δευτερόλεπτα/πλαίσιο (8x ταχύτερος από τον αλγόριθμο *L-SEABI*).

Η Εικόνα 5.9 συνοψίζει όλα τα αποτελέσματα του χρόνου εκτέλεσης, παρουσιάζοντας χρονικά την αποδεκίαση ως τμήμα της κωδικοποίησης και την υπερ-δειγματοληψία ως τμήμα της αποκωδικοποίησης. Είναι προφανές ότι το προτεινόμενο σχήμα ευνοείται ξεκάθαρα από χαμηλής πολυπλοκότητας μεθόδους ΥΑ, όπως είναι οι προτεινόμενες *L-SEABI* και *SIL-SEABI*, οι οποίες μειώνουν το συνολικό χρόνο κωδικοαποκωδικοποίησης έως και κατά 71%. Οι υπόλοιποι μέθοδοι περιορίζουν την αποτελεσματικότητα του σχήματος αφού καταλήγουν σε αύξηση του χρόνου εκτέλεσης: τόσο η μέθοδος *NLIBP*, όσο και η *ANR* χρειάζονται σχεδόν τον ίδιο χρόνο υπολογισμού με αυτόν της συμβατικής κωδικοποίησης μέσω HEVC και 5 φορές περισσότερο από αυτόν της συμβατικής κωδικοποίησης μέσω AVC. Αναφορικά με τις απαιτήσεις σε μνήμη, το προτεινόμενο σχήμα ελαττώνει το αποτύπωμα μνήμης του συμβατικού σχήματος κατά περίπου 2/3 στην περίπτωση του H.264 (από 1.1 σε 0.4 GB) και κατά 3/4 στην περίπτωση του H.265

ΚΕΦΑΛΑΙΟ 5 5.3. ΥΑ μέσω *L-SEABI* στο προτεινόμενο σχήμα συμπίεσης



Εικόνα 5.9: Πολυπλοκότητα συμβατικής και προτεινόμενης συμπίεσης.

(από 2.4 σε 0.6 GB). Επιπρόσθετα, ο αλγόριθμος *L-SEABI* διατηρεί τις χαμηλές απαιτήσεις μνήμης και στον αποκωδικοποιητή, εφόσον χρειάζεται περίπου 35 MB για την υπερ-δειγματοληψία σε 1080p και 95 MB για την υπερ-δειγματοληψία σε μεγέθη 2160p.

ΚΕΦΑΛΑΙΟ 5 5.3. ΥΑ μέσω *L-SEABI* στο προτεινόμενο σχήμα συμπίεσης

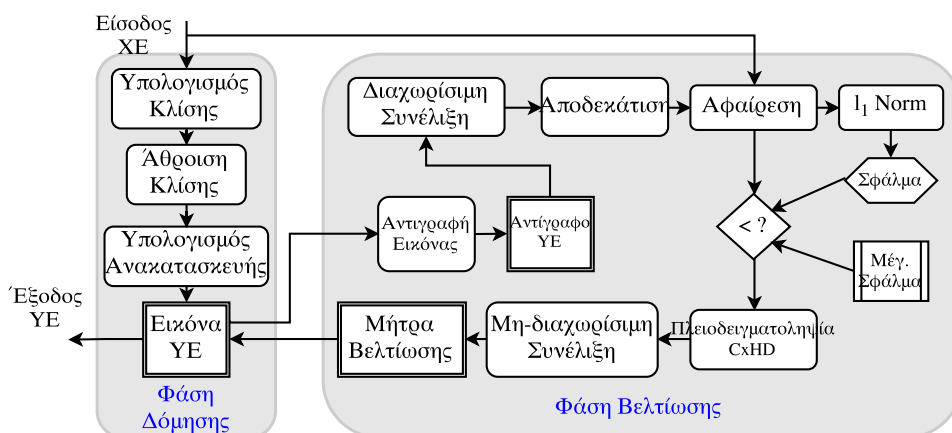
ΚΕΦΑΛΑΙΟ 6

Τεχνικές Παραλληλοποίησης και Υλοποίησης σε Υλικό

Το παρόν κεφάλαιο στοχεύει στη βελτίωση των παρεχόμενων λύσεων σε προβλήματα επεξεργασίας εικόνας, παρουσιάζοντας τεχνικές παραλληλοποίησης για τον αλγόριθμο ΥΑ *SIL-SEABI* [31]. Υπενθυμίζουμε ότι ο αλγόριθμος αυτός σχεδιάστηκε ως μία αποτελεσματική αυτόνομη λύση ΥΑ αλλά και ως τεχνική υπερ-δειγματοληψίας εντός ενός σχήματος συμπίεσης μειωμένης πολυπλοκότητας.

Η μελέτη μας προσανατολίζεται σε δύο κατευθύνσεις. Πρώτον, εξετάζονται τεχνικές παραλληλοποίησης του αλγορίθμου σε μονάδες επεξεργασίας γραφικών γενικού σκοπού (MEG-ΓΣ, General Purpose Graphics Processing Units, GP-GPUs). Για το σκοπό αυτό, αναπτύσσουμε ένα παραμετρικό μοντέλο επιτάχυνσης, το οποίο καθιστά δυνατή την κατατομή του χρόνου εκτέλεσης των επιμέρους διεργασιών του αλγορίθμου. Εν συνεχεία, βελτιστοποιούμε ενδελεχώς το αρχικό μοντέλο της GPU, προκειμένου να επιτύχουμε επεξεργασία πραγματικού χρόνου μέσω τεχνικών βελτιώσεων σε υψηλό και χαμηλό υπολογιστικά επίπεδο. Οι επιμέρους διεργασίες του μοντέλου αναλύονται, βελτιστοποιούνται και τελικώς η απόδοση του μοντέλου αξιολογείται σε τρεις διαδοχικές γενιές MEG με την τελευταία έκδοση της Διεπαφής Προγραμματισμού Εφαρμογής CUDA (ΔΠΕ - Application Programming Interface, API). Τέλος, αποτιμάται η απόδοση του μοντέλου του αλγορίθμου σε σχέση με υλοποιήσεις άλλων μεθόδων υπερ-δειγματοληψίας. Δεύτερον, η μελέτη παραλληλοποίησης του αλγορίθμου εστιάζει στην εφαρμογή σε συστοιχίες πυλών προγραμματίσιμες στο πεδίο (Field Programmable Gate Arrays, FPGAs). Αναλυτικότερα, σχεδιάστηκε μία παραμετρική αρχιτεκτονική παράλληλης επεξεργασίας σε FPGAs η οποία, εκμεταλλευόμενη τη χαμηλή πολυπλοκότητα του αλγορίθμου *SIL-SEABI*, επιτρέπει την επεξεργασία πραγματικού χρόνου από διατάξεις FPGA ποικίλων κατηγοριών.

Συνολικά, απώτερος στόχος της μελέτης μας είναι να καταλήξουμε σε τεχνικές οι οποίες ξεπερνούν τους περιορισμούς της υποκείμενης πλατφόρμας υλοποίησης και τελικά, να παρουσιαστεί μία περιεκτική αξιολόγηση επιδόσεων. Εξ όσων γνωρίζουμε, η διενεργηθείσα αξιολόγηση συμβάλλει στην παρούσα βιβλιογραφία εξετάζοντας από κοινού, σε τρεις διαφορετικές πλατφόρμες, τέσσερις παραμέτρους: α) την απόδοση, β) το πλήθος των δεδομένων εισόδου, γ) την κατανάλωση



Εικόνα 6.1: Οι διακριτές διεργασίες του αλγορίθμου *L-SEABI*.

ενέργειας και δ) το κόστος απόκτησης της πλατφόρμας σε μονάδες επεξεργασίας γραφικών, FPGAs, αλλά και σε πολυπύρηνους επεξεργαστές γενικού σκοπού. Τέλος, για λόγους συνέπειας, εφαρμόζουμε ολόκληρη τη διαδικασία σχεδίασης/βελτιστοποίησης/αξιολόγησης σε έναν υπολογιστικά απαιτητικό αλγόριθμο ανομοιότητας (Disparity).

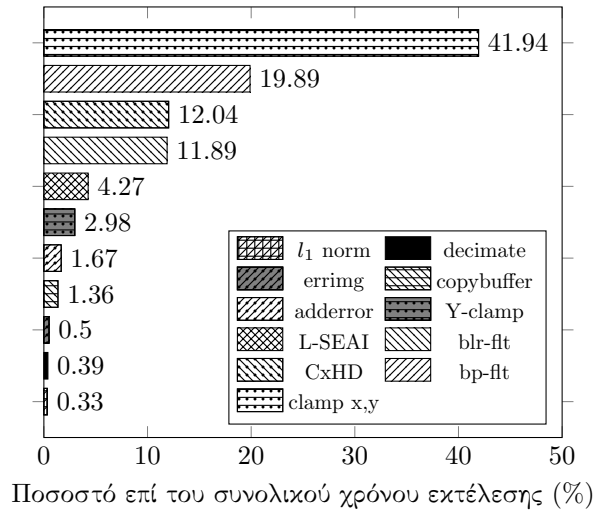
6.1 Αλγόριθμος SIL-SEABI: επιτάχυνση σε ΜΕΓ

6.1.1 Στρατηγική υλοποίησης: από τη CPU στη ΜΕΓ

Στην παρούσα ενότητα περιγράφουμε τα βήματα τα οποία μας οδήγησαν στην υλοποίηση του αλγορίθμου *SIL-SEABI* σε επεξεργαστές γραφικών. Η εικόνα 6.1 απεικονίζει μία επισκόπηση των δύο φάσεων και των διάκριτων διαδικασιών οι οποίες απαρτίζουν τον αλγόριθμο *L-SEABI*. Όλες οι διεργασίες εκτελούνται σειριακά, ξεκινώντας από τον υπολογισμό και την άθροιση των κλίσεων για κάθε εικονοστοιχείο της αρχικής εικόνας. Το αποτέλεσμα της φάσης δόμησης του *L-SEABI* αντιγράφεται σε μία προσωρινή μνήμη και εν συνεχεία, η εικόνα βελτιώνεται μέσα από μία επαναληπτική και προσαρμοστική διαδικασία, εξαρτώμενη από το επιλεγμένο κριτήριο ελαχιστοποίησης σφάλματος.

Προκειμένου να διαμορφώσουμε τη στρατηγική παραλληλοποίησης στη ΜΕΓ, αναλύουμε τη μονονηματική εκτέλεση του αλγορίθμου *SIL-SEABI* μέσω του αναλυτή *oprofile* [52] (το εκτελέσιμο μεταγλωττίστηκε με την GCC 4.9.2). Η αποτίμηση της απόδοσης διενεργείται σε έναν επεξεργαστή Intel Core i5-3470 για 100 εκτελέσεις σε τουλάχιστον 50 διαδοχικά πλαίσια για κάθε εικονοροή από το σύνολο των εξεταζομένων. Τέλος, υπολογίζεται ο μέσος όρος των αποτελεσμάτων σε όλες τις αναλύσεις. Τα παραπάνω φαίνονται στο γράφημα 6.2, απ' όπου προκύπτει ότι περίπου 42% του συνολικού χρόνου εκτέλεσης αφορά στη διεργασία *Clamp x,y* η οποία είναι υπεύθυνη για το πάκτωμα (clamping) των συντεταγμένων των εικονοστοιχείων στα όρια της εικόνας. Το πάκτωμα

SIL-SEABI - Επεξεργαστής Intel i5-3470, Αναλυτής οprofile



Εικόνα 6.2: Κατατομή χρόνου διεργασιών του αλγορίθμου *SIL-SEABI* για μονομηματική εκτέλεση σε επεξεργαστή γενικού σκοπού.

της τιμής της φωτεινότητας στο διάστημα $[0, 255]$ απαιτεί περίπου 3% του συνολικού χρόνου (διεργασία *Y-Clamp*). Το μη-διαχωρίσιμο φίλτρο οπισθοπροβολής 5×5 εικονοστοιχείων (διεργασία *bp-flt*) αντιστοιχεί στο 20% του συνολικού χρόνου, δηλαδή περίπου όσο ο συνδυασμός φάσης δόμησης και υπερ-δειγματοληψίας *CxHD* (12.04% και 4.27% αντίστοιχα). Από την άλλη πλευρά, η διαχωρίσιμη συνέλιξη δεσμεύει 12% του συνολικού χρόνου (διεργασία *Blr-flt*). Έπειτα, η βελτίωση της εικόνας ΥΕ μέσω προσθήκης του οπισθοπροβεβλημένου σφάλματος χρειάζεται 1.36% του συνολικού χρόνου (διεργασία *adderror* στην Εικ. 6.2) η δημιουργία αντιγράφου της ανακατασκευής κατά τη φάση βελτίωσης (διεργασία *copybuffer*) περίπου 1.25% και οι υπόλοιπες διεργασίες αθροίζονται στο 4%. Βάσει των ανωτέρω αποτελεσμάτων κατατομής, θα περιγράψουμε αρχικά τις λεπτομέρειες του γενικευμένου GPU μοντέλου του *L-SEABI*, κατόπιν θα παραθέσουμε τις βελτιώσεις οι οποίες οδηγούν στην προτεινόμενη υλοποίηση του αλγορίθμου *SIL-SEABI* και τέλος θα πραγματοποιήσουμε αποτίμηση των επιδόσεων.

6.1.2 Λεπτομέρειες υλοποίησης σε ΜΕΓ και βελτιστοποιήσεις

Σε αυτή την υποενότητα παρατίθενται α) οι λεπτομέρειες της γενικευμένης υλοποίησης του αλγορίθμου *L-SEABI* σε GPU και β) τα στάδια των βελτιστοποιήσεων τα οποία οδηγούν τελικώς στην προτεινόμενη υλοποίηση του αλγορίθμου *SIL-SEABI*. Το προγραμματιστικό μοντέλο της CUDA επιτρέπει την ενοποιημένη χρήση μαζικά πολυμηματικών επεξεργαστών τόσο για εφαρμογές γραφικών, όσο και για παράλληλους υπολογισμούς γενικού σκοπού. Οι συμβατές με την CUDA συσκευές αποτελούνται από κλιμακοθετήσιμες διατάξεις Προσθηκευματι-

κών Πολυεπεξεργαστών (Streaming Multiprocessors, SMs) στους οποίους ο προγραμματιστής καταναίμει παράλληλα νήματα μέσω κώδικα υψηλού επιπέδου ο οποίος ονομάζεται πυρήνας GPU (GPU kernel) [60]. Η ΔΠΕ της CUDA ορίζει επίσης μία ενδιάμεση συμβολογλώσσα, η οποία ονομάζεται Ιδεατή Μηχανή και Αρχιτεκτονική Συνόλου Εντολών Παράλληλης Εκτέλεσης Νημάτων (Parallel Thread Execution Instruction Set Architecture, PTX ISA) [63]. Ο μεταγλωττιστής της nVidia μεταφράζει τον κώδικα CUDA σε εντολές PTX και εν συνεχεία ο οδηγός γραφικών μετατρέπει τις εντολές PTX σε κώδικα εκτελέσιμο στον γραφικό επεξεργαστή. Για όλα τα πειράματα αποτίμησης που θα ακολουθήσουν, οι δοκιμές του κώδικα GPU εκτελούνται σε τρεις διατάξεις οι οποίες αντιπροσωπεύουν διαδοχικές μείζονες εκδόσεις υπολογιστικών/αρχιτεκτονικών δυνατοτήτων (compute capability). Αναλυτικότερα, οι δοκιμές εκτελούνται στη διάταξη *GeForce GTX 550 Ti* (μικροαρχιτεκτονική Fermi) υπολογιστικών δυνατοτήτων 2.1, τη *GeForce GTX 670* (μικροαρχιτεκτονική Kepler) υπολογιστικών δυνατοτήτων 3.0 και τη *GeForce GTX 960* (μικροαρχιτεκτονική Maxwell) υπολογιστικών δυνατοτήτων 5.2. Το πλήθος των πυρήνων CUDA κάθε συσκευής είναι αντιστοίχως 192, 1344 και 1024. Να σημειωθεί εδώ, ότι η έκδοση των υπολογιστικών δυνατοτήτων κωδικοποιείται έως αυτή τη στιγμή με δύο αριθμητικά ψηφία στη μορφή M.E, από τα οποία το M ορίζει τη μείζονα έκδοση και το E την ελάσσονα. Επίσης, η μετάβαση από τη μείζονα έκδοση 3 έγινε κατευθείαν στη μείζονα έκδοση 5. Κατά την αποτίμησή μας θα παρουσιάσουμε αρχικά μία αδρομερή κατατομή του χρόνου υπολογισμού, για την εκτέλεση ενός βρόχου επανάληψης της γενικευμένης υλοποίησης του αλγορίθμου *L-SEABI*. Βασιζόμενοι στην έκδοση 7.0 της ΔΠΕ (CUDA 7.0) θα συγκρίνουμε με τα αποτελέσματα της Εικ. 6.2 και κατόπιν θα διαμορφώσουμε το μοντέλο επιτάχυνσης προκειμένου να υλοποιεί τον αλγόριθμο *SIL-SEABI*. Θα εστιάσουμε στη γενική ροή των δεδομένων αλλά και τις λεπτομέρειες κάθε πυρήνα GPU, έχοντας ως βάση των βελτιστοποιήσεων το γενικευμένο μοντέλο υλοποίησης του *L-SEABI*. Στόχος μας είναι η τελική υλοποίηση να μη βασίζεται σε συγκεκριμένες - εμπορικές ή μη - βιβλιοθήκες, ώστε να εκτελείται και να είναι αποδοτική ανεξαρτήτως της υποκείμενης αρχιτεκτονικής.

Όπως αναφέρεται στο [75] η χρήση μνήμης κλειδωμένων σελίδων (page-locked μνήμη στη RAM του συστήματος που φιλοξενεί τη GPU) σχεδόν διπλασιάζει τη ρυθμοαπόδοση εν συγκρίσει με τη σελιδοποιημένη μνήμη. Συνεπώς, ξεκινούμε την εκτέλεση του Host κώδικα δεσμεύοντας μνήμη κλειδωμένων σελίδων για τις εικόνες XE και YE μέσω της συνάρτησης *cudaHostAlloc*. Σε περιπτώσεις όπου οι σχηματομορφές προσπέλασης μνήμης παρουσιάζουν χωρική τοπικότητα - όπως στην περίπτωση μας καθ'όλη τη διάρκεια εκτέλεσης του *L-SEABI* - προτιμάται η χρήση μνήμης υφών (texture memory) καθώς τμήματα της μνήμης αυτής αποθηκεύονται αυτόματα και σε λανθάνουσα μνήμη. Συν τοις άλλοις, σε αυτού του τύπου τη μνήμη η διαχείριση των συντεταγμένων εκτός των ορίων της εικόνας εκτελείται χωρίς την παρέμβαση του χρήστη - άρα εξαλείφεται η ανάγκη μιας συνάρτησης αντίστοιχης με την *clamp(x,y)* (Εικ. 6.2). Ως συνέπεια της χρήσης μνήμης υφών, στο μοντέλο μας αντιγράφουμε τις εικόνες από τη μνήμη κλειδωμένων σελίδων στη GPU μέσω της συνάρτησης *cudaMemcpy2D*, μέσω της οποίας

```

Είσοδος: LR Image , max_iterations , max_error
Έξοδος: HR Image
1: gradientCompute $\kappa$ 
2: reduce $\kappa_1 \Rightarrow$  rdc_res $_1$ 
3: lseai $\kappa$ 
for (int i=0; i<max_iterations; i++)
{
4:  cudaMemcpy2D (Device $\rightarrow$ Device)
5:  convCols $\kappa$ 
6:  convRows $\kappa$ 
7:  decimate $\kappa$ 
8:  errimg $\kappa$ 
9:  abs $\kappa$ 
10: reduce $\kappa_2 \Rightarrow$  rdc_res $_2$ 
if (rdc_res $_2$  < max_error)
{
11:     CxHD $\kappa$ 
12:     convNoSep $\kappa$ 
13:     adderror $\kappa$ 
}
else break;
}

```

Εικόνα 6.3: Ψευδοκώδικας υλοποίησης *L-SEABI* στη GPU: ιεραρχία πυρήνων.

χειριζόμαστε και τις απαιτήσεις ευθυγράμμισης της μνήμης υφών σε bytes.

Για λόγους απόδοσης, αποτελεί συνήθη πρακτική ο αριθμός των νημάτων σε ένα block να είναι ακέραιο πολλαπλάσιο του μεγέθους ενός warp (32 νήματα, τα οποία εκτελούνται σε ένα βήμα κλειδώματος - lockstep). Σύμφωνα με τα αποτελέσματα της έρευνας στο [95] ο διαχωρισμός της εικόνας σε διαστάσεις πλακίδια 32×4 δειγμάτων παρέχει τις καλύτερες δυνατές επιδόσεις σε διατάξεις GPU υπολογιστικών δυνατοτήτων 1.3 ή υψηλότερων. Βάσει των ανωτέρω αλλά και των δικών μας δοκιμών ακολουθούμε τον παραπάνω διαχωρισμό στις *GTX 670*, *GTX 960* και αυξάνουμε το μέγεθος του block σε 32×8 νήματα στην *GTX 550Ti*, προσαρμόζοντας κατάλληλα το μέγεθος του grid στις διαστάσεις της εικόνας.

Στην Εικόνα 6.3 παρατίθεται ο ψευδοκώδικας της επεξεργασίας στη GPU για την προτεινόμενη υλοποίηση του *L-SEABI*, υποσημειώνοντας τους πυρήνες της GPU με το δείκτη κ . Στον πυρήνα *gradientCompute κ* κάθε νήμα υπολογίζει την κλίση της φωτεινότητας σε δύο διαστάσεις και παράλληλα αποθηκεύει τα αποτελέσματα των αθροισμάτων των τετραγώνων των κλίσεων σε μία μνήμη $W \times H$ δειγμάτων. Κατόπιν, για τον υπολογισμό της μέσης διακύμανσης χρησιμοποιούμε το σχεδιασμένο από την nVidia πυρήνα παράλληλης περιστολής [35], ο οποίος περιλαμβάνει ξεδίπλωμα διεργασιών σε βρόχους και αθροίσεις πολλαπλών τιμών ανά

νήμα (*reduce_κ* στην Εικ. 6.3). Ο πυρήνας *lsea_{ik}* περιλαμβάνει τον υπολογισμό της αρχικής ανακατασκευασμένης εικόνας βάσει του αποτελέσματος *rdc_res₁* του πυρήνα *gradientCompute_κ* και η κατατομή εκτέλεσής του χαρακτηρίζεται από ισορροπημένο κόστος μεταξύ αριθμητικών υπολογισμών και προσπελάσεων στη μνήμη. Κάθε νήμα α) μεταφέρει 16 γειτονικά δείγματα ανά εικονοστοιχείο από τη μνήμη υφών σε καταχωρητές, β) κατηγοριοποιεί τα εικονοστοιχεία βάσει του κατωφλίου και γ) υπολογίζει τη φωτεινότητα για τις οριζόντιες κατακόρυφες και διαγώνιες θέσεις του υπερ-δειγματοληπτημένου πλέγματος. Να σημειωθεί εδώ ότι για να αποφευχθεί η απόκλιση των νημάτων κατά τον υπολογισμό των συντεταγμένων στο πλέγμα ΥΕ, εκτελούμε τους πυρήνες στη GPU βάσει του σχήματος διαχωρισμού της εικόνας ΧΕ. Εν συνεχεία, οι θέσεις των εικονοστοιχείων ΥΕ υπολογίζονται ως ακολούθως:

$$\begin{aligned}
 I_{offset} &= [(ty * 2) * (W * 2)] + (tx * 2) \\
 HH_{offset} &= [(ty * 2) * (W * 2)] + [(tx * 2) + 1] \\
 HV_{offset} &= \{[(ty * 2) + 1] * (W * 2)\} + (tx * 2) \\
 HD_{offset} &= \{(ty * 2) + 1\} * (W * 2) + [(tx * 2) + 1]
 \end{aligned} \tag{6.1}$$

όπου tx , ty είναι οι δισδιάστατες συντεταγμένες των νημάτων εκπεφρασμένες σε ενιαίο μονοδιάστατο χώρο και W είναι είτε το πλάτος της εικόνας ΧΕ σε εικονοστοιχεία είτε το βήμα της μνήμης σε bytes (κατά τις περιπτώσεις αποθήκευσης της εικόνας ΥΕ σε συνηθισμένη μνήμη ή μνήμη υφών αντιστοίχως). Παρατηρούμε επίσης ότι η προσέγγισή μας ουσιαστικά αυξάνει τον παραλληλισμό επιπέδου εντολής (Instruction-Level Parallelism, ILP) από 1 σε 4 καθώς κάθε νήμα υπολογίζει και αποθηκεύει 4 τιμές εικονοστοιχείων στο υπερ-δειγματοληπτημένο πλέγμα.

Η επαναληπτική βελτίωση της εικόνας (γραμμές 4–13) ξεκινά αντιστοίχως την πρώτη/υπόλοιπες επαναλήψεις αντιγράφοντας την ανακατασκευασμένη/βελτιωμένη εικόνα σε έναν ενδιάμεσο χώρο μνήμης δεσμευμένο σε υφή. Η αντιγραφή αυτή λαμβάνει χώρο εντός της διάταξης GPU και προηγείται των δύο πυρήνων διαχωρίσιμης συνέλιξης (*convCols_κ*, *convRows_κ*), οι οποίοι επίσης βασίζονται σε μνήμη υφών και ενδιάμεσους καταχωρητές. Αναλυτικότερα, κάθε νήμα μεταφέρει 5 τιμές φωτεινότητας ανά εικονοστοιχείο από τη μνήμη ΥΕ σε καταχωρητές της GPU, υπολογίζει το άθροισμα της συνέλιξης και αντιγράφει το αποτέλεσμα από καταχωρητές στην κεντρική μνήμη της GPU. Να σημειωθεί εδώ ότι όλες οι διεργασίες φιλτραρίσματος στους σχεδιασμένους πυρήνες βασίζονται σε ξεδίπλωμα των σχετικών διεργασιών οι οποίες απαρτίζουν τους βρόχους.

Έπειτα η εικόνα αποδεκατίζεται μέσω του πυρήνα *decimate_κ*, ο οποίος χρησιμοποιεί τη διαμόρφωση νημάτων της εικόνας ΧΕ και απλώς αντιγράφει δεδομένα από τα εικονοστοιχεία της μνήμης ΥΕ με συντεταγμένες $(tx*2, ty*2)$, στη μνήμη ΧΕ στις θέσεις (tx, ty) . Οι πυρήνες *cxHD_κ* και *convNoSep_κ* ακολουθούν τις ίδιες σχεδιαστικές επιλογές με τον πυρήνα του αλγορίθμου *L-SEAI*: βασίζονται σε αναγνώσεις από μνήμη υφών και αποθηκεύουν τα ενδιάμεσα αποτελέσματα σε καταχωρητές. Συγκεκριμένα, τα νήματα εντός του πυρήνα *cxHD_κ* διαβάζουν 12 εικονοστοιχεία ανά δείγμα εισόδου και τα νήματα του πυρήνα *convNoSep_κ*

$convCols_K$, $convRows_K$ με συμβολισμό $blr-flt$ αντίστοιχα στην Εικ. 6.4. Επίσης, παρατηρούμε μία εμφανή διαφοροποίηση στο ποσοστό του χρόνου εκτέλεσης μεταξύ των δύο διατάξεων GPU. Από τη στιγμή που ο πυρήνας $convNoSep_K$ απαιτεί 25 αναγνώσεις εικονοστοιχείων ανά δείγμα εισόδου, η απόδοσή του περιορίζεται από τη διεκπεραιωτικότητα της μνήμης υφών η οποία φτάνει το όριο των 232.7 GB/s στη *GTX 550 Ti* (680.7 GB/s στη *GTX 960*). Η χρήση πόρων εμφανίζεται πιο ισορροπημένη όταν εξετάζουμε τους πυρήνες οι οποίοι εκτελούν τη διαχωρίσιμη συνέλιξη. Στην περίπτωση του $convRows_K$ η απόδοση δεσμεύεται από την ταχύτητα των υπολογιστικών μονάδων καθώς όλες οι αναγνώσεις αναφέρονται σε διαδοχικές διευθύνσεις μνήμης, αντίθετα με τον πυρήνα $convCols_K$ η απόδοση του οποίου δεσμεύεται από το υποσύστημα μνήμης. Σημαντικό είναι επίσης και το κόστος των συναλλαγών μνήμης εκτός πυρήνων, καθώς αντιστοιχούν σε περίπου 26% του συνολικού χρόνου. Ειδικότερα, οι συναλλαγές μνήμης από και προς τη GPU σχετίζονται με τη μεταφορά των εικόνων XE/YE και είναι υποχρεωτικές. Πέρα λοιπόν από την ήδη υιοθετημένη προσέγγιση της μνήμης κλειδωμένων σελίδων δεν υπάρχει περαιτέρω περιθώριο βελτίωσης των συναλλαγών αυτών. Εξάλλου, κύριος στόχος μας είναι η επιτάχυνση των διεργασιών επί της GPU. Να σημειώσουμε εδώ ότι υπάρχει επιπλέον ένα αμελητέο κόστος σε χρόνο, σχετιζόμενο με τη μεταφορά των δύο ακέραιων αποτελεσμάτων των πυρήνων $reduce_K$ από τη GPU. Ωστόσο, οι συναλλαγές μνήμης εντός της GPU (με συμβολισμό D2D στην Εικ. 6.4) δύνανται να ελαχιστοποιηθούν καθώς αφορούν τη μεταφορά δεδομένων σε μνήμη υφών.

Στην επόμενη ενότητα περιγράφονται οι προτεινόμενες βελτιστοποιήσεις της υλοποίησης σε GPU και κατηγοριοποιούνται βάσει της εμβέλειάς τους ως *αλγοριθμικού επιπέδου* και *επιπέδου ΔΠΕ*. Εκτός αυτού, θα αποτιμηθεί ο απότοκος αυτών των βελτιστοποιήσεων συγκρίνοντας το κέρδος σε απόδοση εν συγκρίσει με την προτεινόμενη γενικευμένη υλοποίηση του *L-SEABI*. Οι τροποποιημένοι πυρήνες της GPU υποσημειώνονται με το γράμμα “m” στην αρχή της συμβολικής τους ονομασίας.

Βελτιστοποιήσεις αλγοριθμικού επιπέδου

Πρώτα απ’ όλα, οι βελτιστοποιήσεις σε επίπεδο αλγορίθμου είναι εφικτές εφόσον στόχος των προτεινόμενων τεχνικών είναι η εκτέλεση μίας επανάληψης του αλγορίθμου *L-SEABI*. Συνεπώς, μπορούμε να θεωρήσουμε εκ των προτέρων ότι η προσαρμοστική οπισθοπροβολή του σφάλματος εκτελείται πάντοτε. Για την προτεινόμενη υλοποίηση του *SIL-SEABI* μπορούμε λοιπόν να παραλείψουμε την εκτέλεση του πυρήνα abs_K στη γραμμή 9 της Εικ. 6.3, καθώς και το δεύτερο πυρήνα συσσώρευσης στη γραμμή 10 - οι οποίοι κατά μέσο όρο αντιστοιχούν στο 3.75% του συνολικού χρόνου εκτέλεσης (διεργασία l^1 norm στην Εικ. 6.4). Επιπρόσθετα, τροποποιούμε τον πυρήνα $cxHD_K$ περιορίζοντας τους γειτονικούς ελέγχους από 8 σε 4 εικονοστοιχεία για κάθε ΗΔ θέση. Καταλήγουμε έτσι στον τροποποιημένο πυρήνα $mcxHD_K$, ο οποίος εφαρμόζει το ίδιο διγραμμικό φίλτρο: η βελτίωση έγκειται στο ότι απαιτεί 2 λιγότερους καταχωρητές ανά νήμα και 12

λιγότερες πράξεις σύγκρισης για την κατηγοριοποίηση των ακμών, σε σχέση με τις 14 του αρχικού αλγορίθμου *CxHD*.

Για τη βελτιστοποίηση της διαχωρίσιμης συνέλιξης, εκμεταλλευόμαστε αρχικά το γεγονός ότι η Γκαουσιανή $\Sigma\Delta\Sigma$ 5×5 δειγμάτων περιέχει συμμετρικούς συντελεστές. Άρα, μπορούμε να ελαττώσουμε τον αριθμό των πολλαπλασιασμών σε 3 ανά εικονοστοιχείο εισόδου γράφοντας το άθροισμα της συνέλιξης ως $p^{out} = h_0 * p_0^{in} + h_1 * p_1^{in} + h_2 * p_2^{in} + h_3 * p_3^{in} + h_4 * p_4^{in} = h_0 * (p_0^{in} + p_4^{in}) + h_1 * (p_1^{in} + p_3^{in}) + h_2 * p_2^{in}$, όπου p^{out} είναι η τελική φωτεινότητα του εικονοστοιχείου, h_n και p_n^{in} , $n \in [0, 4]$ οι συντελεστές του φίλτρου και οι αρχικές τιμές φωτεινότητας αντιστοίχως. Καλούμε επίσης τον αναγνώστη να παρατηρήσει ότι στην Εικ. 6.3 ο πυρήνας *convRowsκ* ακολουθείται πάντοτε από τον πυρήνα *decimateκ*. Η σύμπτυξη αυτών των δύο διαδικασιών στον πυρήνα *convRowsdecκ* παρέχει τα εξής δύο οφέλη: α) παρακάμπτει το επίβαρο κόστος δημιουργίας και εκτέλεσης ενός διάκριτου πυρήνα αποδεκάτισης και β) οδηγεί στην ελάττωση του συνολικού αριθμού υπολογισμών και αναγνώσεων σχετιζόμενων με τον πυρήνα *convRowsκ* καθώς λόγω της άμεσης αποδεκάτισης, η συνέλιξη χρειάζεται ουσιαστικά μόνο τα άρτια εικονοστοιχεία. Προκειμένου να αποφευχθεί η εμφάνιση νηματικών αποκλίσεων ακολουθούμε παρόμοια προσέγγιση με αυτή της Εξίσωσης 6.1: υιοθετούμε τη νηματική διαμέριση της εικόνας ΧΕ και διαβάζουμε τα άρτια δείγματα από τη μνήμη υφών ως ακολούθως:

$$p_n^{in} = tex2D(texref, tx*2 + n - 2, ty*2), n \in [0, 4] \quad (6.2)$$

όπου *tex2D* είναι η συνάρτηση αναζήτησης από μνήμη υφών της ΔΠΕ CUDA και *texref* είναι η δεσμευμένη σε υφές μνήμη στην οποία περιέχεται η φιλτραρισμένη ανά στήλες ανακατασκευασμένη εικόνα ΥΕ.

Βελτιστοποιήσεις βασισμένες στην πλατφόρμα υλοποίησης

Για την εφαρμογή περισσότερο εξειδικευμένων βελτιστοποιήσεων θα μελετήσουμε την υλοποίησή μας τόσο σε υψηλό επίπεδο (κώδικας CUDA C), όσο και σε χαμηλό (κώδικας αρχιτεκτονικής συνόλου εντολών PTX), με σκοπό να προτείνουμε συγκεκριμένες τροποποιήσεις. Η κατατομή του χρόνου εκτέλεσης στη GPU (Εικ. 6.4) έδειξε ότι το μεγαλύτερο μέρος του επεξεργαστικού χρόνου δαπανάται στη μη διαχωρίσιμη συνέλιξη κατά τη φάση βελτίωσης του αλγορίθμου *L-SEABI*. Από τη στιγμή που οι επιδόσεις της συνέλιξης περιορίζονται από το ζωνικό εύρος της μνήμης, η απόδοση του πυρήνα *convNoSepκ* μπορεί να βελτιωθεί μέσω διαμοιρασμένης μνήμης. Για το σκοπό αυτό βασιζόμαστε στο σχήμα χρήσης που διατύπωσαν οι Eklund κ.ά. στο [25] εφόσον είναι συμβατό με όλες τις αρχιτεκτονικές GPU που εξετάζουμε: νηματικές ομάδες αποτελούμενες από 32×32 νήματα μεταφέρουν τμήματα εικόνας 96×64 εικονοστοιχείων προς τη διαμοιρασμένη μνήμη προκειμένου να εφαρμόσουν δισδιάστατη συνέλιξη. Να σημειωθεί εδώ ότι η *GTX 960* έχει 33% περισσότερη διαμοιρασμένη μνήμη, γεγονός το οποίο μας επιτρέπει να τροποποιήσουμε το ανωτέρω σχήμα αυξάνοντας το μέγεθος του παραθύρου σε 128×64 εικονοστοιχεία.

Πίνακας 6.1: Τύπος εντολής PTX για τη διεργασία της διαχωρίσιμης συνέλιξης, αριθμός εκτελέσεων και ρυθμοαπόδοση ανά αριθμητική/επεξεργαστή γραφικών.

<i>blr-fft</i> Αριθμ. Τύπος Εντολής	Εκτελέσεις	εντολές/κύκλο/ΠΕ	
ακεραίων ■ κινητής υποδιαστολής		550Ti	670
and.b32 ■ and.b16	5	48 ■ 48	160 ■ 160
add.s32 ■ add.ftz.f32	2	48 ■ 48	160 ■ 192
mul.lo.s32 ■ mul.ftz.f32	1	16 ■ 48	32 ■ 192
mad.lo.s32 ■ fma.rm.ftz.f32	2	16 ■ 48	32 ■ 192
shr.u32 ■ cvt.rzi.ftz.u32.f32	1	16 ■ 16	32 ■ 32
- ■ cvt.u16.u32	5	- ■ 16	- ■ 32
- ■ cvt.rm.f32.u16	5	- ■ 16	- ■ 128

Η αναλυτική κατατομή του χρόνου εκτέλεσης των προτεινόμενων πυρήνων διαχωρίσιμης συνέλιξης μας έδειξε ότι η απόδοσή τους περιορίζεται από την ταχύτητα των επεξεργαστικών μονάδων της GPU. Από σχετική έρευνα [91] προκύπτει ότι η αύξηση του επιπέδου παραλληλισμού εντολής (Instruction Level Parallelism, ILP) δύναται να οδηγήσει σε βελτίωση επιδόσεων ακόμη και όταν γεννώνται λιγότερα νήματα. Στη δική μας περίπτωση, τροποποιούμε τους πυρήνες *convRowsK* και *convColsK* έτσι ώστε κάθε νήμα να μεταφέρει ένα σύνολο 5 εικονοστοιχείων από 4 διαδοχικές νηματικές ομάδες, να υπολογίζει 4 αθροίσματα συνέλιξης και τελικώς να αποθηκεύει 4 αποτελέσματα. Αυτό αυξάνει τον αριθμό των απαιτούμενων καταχωρητών ανά νήμα από 12 σε 28 στην *GTX 550Ti*, από 14 σε 31 στην *GTX 670* και από 13 σε 31 στη *GTX 960*. Στους τροποποιημένους πυρήνες προσεγγίζεται το όριο του εύρους ζώνης της μνήμης υφών, συνεπώς για περαιτέρω αύξηση των επιδόσεων θα στραφούμε στη χρήση της διαμοιρασμένης μνήμης τροποποιώντας το παράδειγμα διαχωρίσιμης συνέλιξης της nVidia [71]. Μετά την επαναδιαμόρφωση των πυρήνων για ILP=4, το εύρος ζώνης της μνήμης υφών υπερβαίνει τα 800GB/s στη *GTX 670* - χωρίς περιορισμό επιδόσεων - και το αντίστοιχο εύρος για το βασισμένο στη Maxwell αρχιτεκτονική επεξεργαστή αγγίζει τα 534GB/s. Να θυμηθούμε εδώ ότι οι λειτουργίες του *L-SEABI* είχαν σχεδιαστεί ώστε να βασίζονται σε αριθμητική ακεραίων με εξαίρεση τη διεργασία διαχωρίσιμης συνέλιξης [31]. Αναλυτικότερα, για να ενισχύσουμε την ποιότητα του αποτελέσματος εφαρμόζουμε συνέλιξη μέσω των σταθερών $\langle h_0, h_1, h_2 \rangle = \langle 0.00257, 0.165795, 0.664904 \rangle$ και μετατρέπουμε το αποτέλεσμα σε ακέραιο αριθμό. Κατά πρόσφορο τρόπο, οι σύγχρονοι επεξεργαστές γραφικών ενσωματώνουν πρωτογενώς μονάδες υπολογισμού κινητής υποδιαστολής η απόδοση των οποίων σε λειτουργίες πολλαπλασιασμού υπερβαίνει εκείνη των πράξεων μεταξύ ακεραίων. Η ρυθμοαπόδοση των συνδυασμένων πολλαπλασιασμών-προσθέσεων μεταξύ 32-bit αριθμών κινητής υποδιαστολής φθάνει τις 48, 192 και 128 πράξεις (operations, ops) ανά κύκλο ρολογιού (clock cycle, cc) ανά πολυεπεξεργαστή (multiprocessor, MP) στις *GTX 550Ti*, *GTX 670* και *GTX 960* αντιστοίχως. Να σημειώσουμε βέβαια ότι απαιτείται ο ρητός ορισμός των συντελεστών ως αριθμών κινητής υποδιαστολής μονής ακρίβειας, ειδάλλως ο μεταγλωττιστής θα θεωρήσει διπλή ακρίβεια η οποία συνεπάγεται σαφώς χαμηλότερη

ρυθμοαπόδοση (4, 8 και 1 ops/cc/MP αντιστοίχως στους ανωτέρω επεξεργαστές γραφικών). Προκειμένου να διατηρήσουμε τη συμβατότητα με τον υπόλοιπο κώδικα (στον οποίο γίνονται υπολογισμοί και αποθηκεύσεις μεταξύ ακεραίων) απαιτούνται πρόσθετες εντολές μετατροπής μεταξύ αναπαραστάσεων. Από τη μελέτη του κώδικα PTX φαίνεται ότι εκτός των εντολών υπολογισμού, η συνέλιξη με αριθμούς κινητής υποδιαστολής χρειάζεται 5 μετατροπές από 32 σε 16-bit μη προσημασμένων, 5 μετατροπές από 16 σε 32-bit κινητής υποδιαστολής και 1 μετατροπή από κινητής υποδιαστολής σε 32-bit μη προσημασμένου. Οι παραχθείσες εντολές της διαχωρίσιμης συνέλιξης για τις περιπτώσεις ακεραίων και κινητής υποδιαστολής απεικονίζονται στον πίνακα 6.1 όπου αντιπαραβάλλονται με την αντίστοιχη σε κάθε περίπτωση ρυθμοαπόδοση. Στους επεξεργαστές αρχιτεκτονικής Fermi και Kepler η ρυθμοαπόδοση του συνδυασμού πολ/μού-άθροισης μεταξύ ακεραίων 32 ψηφίων είναι 16 και 32 ops/cc/MP αντίστοιχα [61]. Για λόγους σύγκρισης με την αντίστοιχη υλοποίηση κινητής υποδιαστολής ο υπολογισμός των Γκαουσιανών συντελεστών γίνεται με τον πολ/μό των αρχικών συντελεστών με το 2^{23} , καταλήγοντας στο φίλτρο $(h_0, h_1, h_2) = (21559, 1390789, 5577619)$. Παρατηρούμε επίσης ότι η πράξη της διαίρεσης μετατρέπεται από το μεταγλωττιστή σε αριθμητική ολίσθηση (*shr.u32*). Ο πίνακας 6.1 δείχνει ότι παρά το γεγονός πως η ρυθμοαπόδοση του πολ/μού κινητής υποδιαστολής είναι εμφανώς υψηλότερη εν συγκρίσει με τις πράξεις ακεραίων, το πλήθος των απαιτούμενων μετατροπών είναι μη αμελητέο άρα ο αντίκτυπός τους στο συνολικό χρόνο εκτέλεσης απαιτεί περαιτέρω διερεύνηση. Η ρυθμοαπόδοση της GTX 960 είναι σε ορισμένες περιπτώσεις χαμηλότερη αυτής της GTX 670 - κάτι που ωστόσο αντισταθμίζεται από τη μειωμένη καθυστέρηση εκτέλεσης εντολών και το μειωμένο αριθμό αδρανών κύκλων της βασισμένης στην αρχιτεκτονική Maxwell MEG.

Οι βελτιστοποιήσεις των διεργασιών υπερ-δειγματοληψίας εστιάζουν στους πυρήνες *lseaiκ* και *cxHDκ* και βασίζονται σε παρατηρήσεις επί του PTX κώδικα. Καλούμε τον αναγνώστη να παρατηρήσει ότι ο *L-SEABI* απαιτεί τον υπολογισμό μίας τετραγωνικής ρίζας για το κατώφλι του χαρακτηρισμού των ακμών [31]. Ο υπολογισμός της ρίζας στην πλατφόρμα CUDA υλοποιείται σε δύο στάδια, με το πρώτο να υπολογίζει τον αντίστροφο της ρίζας και το δεύτερο τον αντίστροφο του αποτελέσματος. Σε κώδικα PTX αυτό μεταφράζεται στην εντολή *sqrt.approx.ftz.f32*. Εξαιτίας της - βασισμένης σε αριθμητική ακεραίων - φύσης του *L-SEABI* και της τιμής του κατωφλίου $T = \sqrt{\frac{TV^{LR}}{WH}}$ όπου $TV^{LR} = rdc_res1$ (Εικ. 6.3) θεωρούμε δύο εναλλακτικούς τρόπους υπολογισμού: α) θέτουμε $T = \frac{TV^{LR}}{WH}$ και κατόπιν υπολογίζουμε το τετράγωνο των διαφορών φωτεινότητας, β) υπολογίζουμε τον αντίστροφο ρίζας $\frac{WH}{TV^{LR}}$. Εκτός των δύο αυτών επιλογών, στην επόμενη ενότητα αποτιμούμε και την απόδοση ενός υπολογισμού κατωφλίου χαμηλότερης ακρίβειας, βασισμένου σε ακέραια διαίρεση. Τέλος, βασισμένοι στον παραγόμενο κώδικα PTX τροποποιούμε όλες τις συγκρίσεις ώστε να έχουν τη μορφή λογικών δηλώσεων και όλες τις συγκρίσεις εντός των πυρήνων *mlseaiκ* και *mcxHDκ* προκειμένου να εξαλειφθούν οι εμφωλεασμένες συνθήκες *if* με τη βοήθεια συγχωνευμένων διακλαδώσεων (branch fusion). Ως παράδειγμα των ανωτέρω, η Εικ. 6.5 παρουσιάζει τον τρόπο μετατροπής του υπολογισμού

***mcxHD_K* Υπολογισμός ΗΟ θέσεων:**

Αρχικές Λογικές Συνθήκες:

- 1: $A > 0$ // συνθήκη #1, κατακόρυφη ακμή
- 2: $A < 0$ // συνθήκη #2, οριζόντια ακμή
- 3: $B > 0$ // συνθήκη #3, NE διεύθυνση
- 4: $B < 0$ // συνθήκη #4, SW διεύθυνση

Υπολογισμός με εμφωλεασμένες συνθήκες if:

- 5: **if** $A > 0$:
- 6: **if** $B > 0$: $p_{HH}^{out} \leftarrow$ δικυβική παρεμβολή #1
- 7: **else if** $B < 0$: $p_{HH}^{out} \leftarrow$ δικυβική παρεμβολή #2
- 8: **else**: $p_{HH}^{out} \leftarrow$ διγραμμική παρεμβολή
- 9: **else if** $A < 0$: $p_{HH}^{out} \leftarrow$ δικυβική παρεμβολή #3
- 10: **else**: $p_{HH}^{out} \leftarrow$ διγραμμική παρεμβολή

Πρόσθετες Λογικές Συνθήκες:

- 11: $A = 0$ // συνθήκη #5, ομοιογένεια
- 12: $B = 0$ // συνθήκη #6, αποκλειστικά κατακόρυφη

Υπολογισμός άνευ νηματικών αποκλίσεων:

- 13: $p_{HH}^{out} \leftarrow$
 (συνθ#1 && συνθ#3) * (δικυβική #1) +
 (συνθ#1 && συνθ#4) * (δικυβική #2) +
 (συνθ#2) * (δικυβική #3) +
 (συνθ#6 | (συνθ#1 && συνθ#5)) * (διγραμμική)

Εικόνα 6.5: Κώδικας υπολογισμού εικονοστοιχείων σε ΗΟ θέσεις εντός του πυρήνα *mcxHD_K*: υπολογισμός χωρίς νηματικές αποκλίσεις.

των ΗΟ εικονοστοιχείων p_{HH}^{out} στον πυρήνα *mcxHD_K* (γραμμές 5-10) σε μία βελτιστοποιημένη έκδοση χωρίς νηματικές αποκλίσεις (γραμμή 13). Παρατηρούμε ότι σε αυτή την περίπτωση θα πρέπει να οριστούν ρητά δύο πρόσθετες λογικές συνθήκες (γραμμές 11-12) για να προκύψει ταυτόσημο αποτέλεσμα. Εκ πρώτης όψεως, φαίνεται ότι μετατρέπουμε τη συνθήκη διακλάδωσης σε ένα άθροισμα γινομένων ακεραίων. Στην πραγματικότητα όμως, μόνο ένα από τα γινόμενα αυτά συμβάλλει στην τιμή του p_{HH}^{out} για κάθε εικονοστοιχείο εισόδου. Επιπρόσθετα, από τη στιγμή που κάθε τελεστέος στα αριστερά των πολ/μών αποτιμάται σε 0 ή 1, μπορούμε να βελτιώσουμε την απόδοση μετατρέποντας κάθε πολ/μό σε δύο δυφιηδόν και μία ακέραια πράξη άθροισης ως εξής: $cmd * flt = [!(cmd) + 1] \& flt$, όπου $cmd \in \{0, 1\}$ είναι η λογική συνθήκη, flt το ακέραιο αποτέλεσμα του φίλτρου, $!$ ο τελεστής δυφιηδόν τελεστής άρνησης (*NOT*) και $\&$ ο δυφιηδόν τελεστής σύζευξης (*AND*). Εν συγκρίσει με τον πολ/μό ακεραίων, αμφότερες οι λειτουργίες αυτές χαρακτηρίζονται από σημαντικά υψηλότερη ρυθμοαπόδοση, ειδικά στους νεότερους επεξεργαστές γραφικών [61].

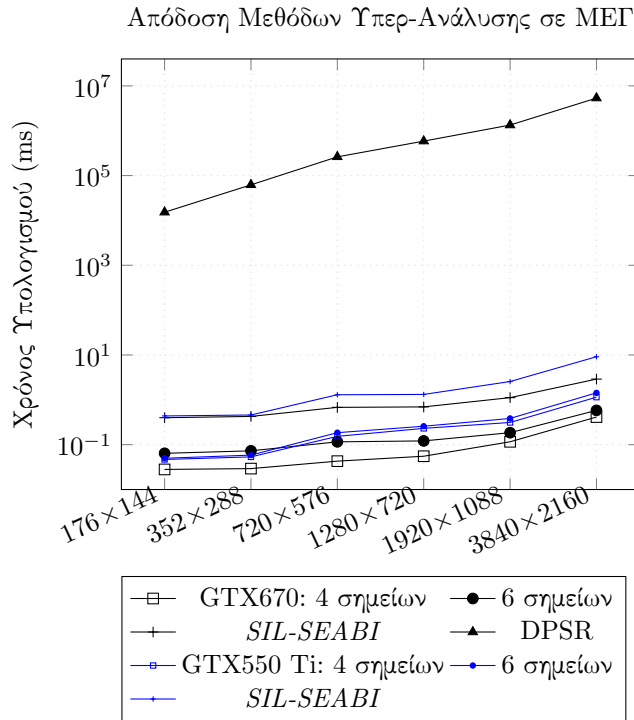
Όσον αφορά τις υπόλοιπες προτεινόμενες βελτιστοποιήσεις, στόχος είναι ο υπολογισμός των απολύτων διαφορών στους πυρήνες *mgradientCompute_K*, *mlseai_K* και *mcxHD_K*. Για τον υπολογισμό της απόλυτης διαφοράς στη φωτεινότητα μεταξύ των εικονοστοιχείων p_0^{in} και p_1^{in} , υιοθετούμε την ενδογενή αριθμητική

επιτάχυνση της τάξης του 1.16x στον πυρήνα *CxHDκ* (cxaug στην Εικ. 6.6). Ο δε συνδυασμός της αποδεκτικής με την διαχωρίσιμη συνέλιξη κατά τον οριζόντιο άξονα βάσει της προτεινόμενης λύσης άνευ νηματικών αποκλίσεων (Εξ. 6.2, defusion στην Εικ. 6.6) καταλήγει σε εμφανέστερη βελτίωση της απόδοσης καθώς η επιτάχυνση του πυρήνα *convRowsκ* ξεπερνά το 2.7x. Ωστόσο, η αναδιάταξη του κώδικα συνέλιξης προκειμένου να εκμεταλλευτούμε τη συμμετρία του Γκαουσιανού πυρήνα δεν παρέχει πρόσθετη επιτάχυνση καθώς ο μεταγλωττιστής παράγει ταυτόσημο κώδικα PTX και στις δύο περιπτώσεις (αναδιάταξης η μη).

Όπως φαίνεται από τα υπόλοιπα ραβδογράμματα, οι βασισμένες στην πλατφόρμα υλοποίησης βελτιστοποιήσεις επιτυγχάνουν τις υψηλότερες τιμές επιτάχυνσης. Αναλυτικότερα, η χρήση της διαμοιρασμένης μνήμης στον πυρήνα μη διαχωρίσιμης συνέλιξης *mconvNoSerκ* (κωδικοποιημένη ως *bp-flt: smem* στην Εικ. 6.6) έχει ιδιαίτερα σημαίνοντα ρόλο, μειώνοντας το χρόνο εκτέλεσης του αρχικού πυρήνα κατά 4.65x στον επεξεργαστή *GTX 670* και πάνω από 6x στους *GTX 550Ti* και *GTX 960* οι οποίοι διέπονται από χαμηλότερη ρυθμοαπόδοση μνήμης υφών. Να σημειωθεί εδώ ότι η αξιοποίηση της αυξημένης διαμοιρασμένης μνήμης του *GTX 960* μέσω της κατάλληλης διαμόρφωσης της μεθόδου στο [25] προκειμένου να επεξεργαστούμε μια περιοχή 124×60 δειγμάτων και συνεπώς να αυξήσουμε τον ΠΕΕ σε 8 από 6, αυξάνει ελαφρώς την αξιοποίηση των υπολογιστικών πόρων κατά 1.5% χωρίς ωστόσο να εμφανίζει σημαντικό κέρδος. Αυτό συμβαίνει εξαιτίας του μικρού μεγέθους του πυρήνα συνέλιξης 5×5 δειγμάτων. Όσον αφορά τη διαχωρίσιμη συνέλιξη βασισμένη σε μνήμες υφών, η αύξηση του ΠΕΕ σε 4 για κάθε νήμα στους πυρήνες *convColsκ* και *convRowsκ* μειώνει το χρόνο τους κατά 1.6x στον *GTX 550Ti* και κατά 1.15x στον *GTX 670* (*blr-flt: ILP4, tex* στην Εικ. 6.6). Η βασισμένη σε μνήμες υφών υλοποίηση επιταχύνεται περαιτέρω αν αξιοποιηθεί η διαμοιρασμένη μνήμη και ταυτόχρονα διπλασιαστεί ο ΠΕΕ (*blr-flt: ILP8, smem* στην Εικ. 6.6, [71]). Εξαιτίας του μικρού μεγέθους του Γκαουσιανού πυρήνα συνέλιξης, της υψηλής ρυθμοαπόδοσης των καταχωρητών των επεξεργαστών γραφικών και της χωρικής τοπικότητας της συνέλιξης, η επιτευχθείσα επιτάχυνση είναι περισσότερο εμφανής στον πυρήνα *convRowsκ* (επιτάχυνση κατά περίπου 2.6x). Αυτό το επίπεδο παραλληλισμού είναι το μέγιστο που μπορεί να επιτευχθεί σε αυτούς τους πυρήνες και αυτούς τους επεξεργαστές γραφικών, καθώς οποιαδήποτε περαιτέρω αύξηση οδηγεί σε καθυστερήσεις αριθμητικών υπολογισμών οφειλόμενες σε χαμηλή αξιοποίηση των νηματικών ομάδων. Η αποτίμηση της χρήσης των πράξεων ακεραίων κατά τη διεργασία φιλτραρίσματος του *SIL-SEABI* επιβεβαιώνει τις παρατηρήσεις του Πίνακα 6.1: η απουσία των πρόσθετων απαιτούμενων πράξεων μετατροπής σε αριθμούς κινητής υποδιαστολής ευνοεί ελαφρώς τη σχετική απόδοση της υλοποίησης ακεραίων κατά 1.05x κατά μέσο όρο. Χαρακτηριστικά εμφανές είναι το αποτέλεσμα το οποίο σχετίζεται με τις μεταφορές δεδομένων εντός του χώρου μνήμης της διάταξης GPU (D2D). Περιορίζοντας τις μεταφορές αυτού του τύπου σε μία (εγγράφοντας απευθείας σε μνήμη υφών ή μέσω της χρήσης διαμοιρασμένης μνήμης) μειώνει τον αντίστοιχο χρόνο μεταφορών κατά περίπου 5x εν συγκρίσει με την αρχική υλοποίηση.

Εν συνεχεία, αποτιμούμε τις προτεινόμενες βελτιστοποιήσεις οι οποίες σχετίζονται με τους πυρήνες *lseaiκ* και *cxHDκ*. Η παράλειψη του υπολογισμού ενός αντιστρόφου ρίζας παρέχει μικρό πλεονέκτημα σε σχέση με τον υπολογισμό του επεξεργαστή γραφικών (*rsqrt/sqrt* στην Εικ. 6.6) κάτι αναμενόμενο εφόσον ο εν λόγω υπολογισμός πραγματοποιείται μόνο μία φορά για την τιμή του κατωφλίου στον πυρήνα *lseaiκ*. Καλούμε επίσης τον αναγνώστη να παρατηρήσει ότι αποφεύγοντας τον υπολογισμό της ρίζας στον τύπο του κατωφλίου, ουσιαστικά αντικαθιστούμε τον αντίστροφο ρίζας με 6 ακέραιες πράξεις πολλαπλασιασμού και συνεπώς ελαττώνουμε τη ρυθμοαπόδοση των επεξεργαστών γραφικών (*rsqrt/nosqrt* στην Εικ. 6.6). Κατά συνέπεια, ο αντίστροφος ρίζας παρέχει τη μέγιστη δυνατή απόδοση κατά τον υπολογισμό του κατωφλίου T στον πυρήνα *lseaiκ*. Μπορούμε επίσης να “ανταλλάξουμε” αριθμητική ακρίβεια με οριακό κέρδος στην απόδοση μέσω της αποκλειστικής χρήσης ακέραιας αριθμητικής κατά τη διαίρεση $\frac{WH}{\sqrt{LR}}$ όπως φαίνεται από τα ραβδογράμματα με κωδικοποίηση *lrrsqrt/rsqrt* στην Εικ. 6.6. Εντός των πυρήνων *lseaiκ* και *cxHDκ* χρησιμοποιούνται εκτεταμένα βελτιστοποιήσεις διακλάδωσης (κωδικοποιημένες ως “συγκρίσεις” στην Εικ. 6.6). Ειδικά όσες εξ αυτών απεικονίζονται στην Εικ. 6.5 παρουσιάζουν μέση επιτάχυνση έως και 1.35x στους προαναφερθέντες πυρήνες. Η αύξηση στην απόδοση είναι εμφανέστερη στον πυρήνα *cxHDκ* σε σχέση με τον πυρήνα *lseaiκ* (1.48x αντί 1.15x κατά μέσο όρο) καθώς ο μεταγλωττιστής παράγει κατηγορηματικές εντολές (predicated instructions) για τον τελευταίο εξαιτίας της απουσίας πολλαπλών/εμφωλεασμένων συνθηκών στον αρχικό αλγόριθμο [31]. Επίσης, το κέρδος είναι πιο έντονο στο χαμηλότερης υπολογιστικής δυνατότητας επεξεργαστή *GTX 550Ti*, καθώς τόσο ο *GTX 670* όσο και ο *GTX 960* μπορούν να χρονοπρογραμματίσουν τέσσερις νηματικές ομάδες και δύο ανεξάρτητες εντολές ανά νηματική ομάδα και ανά κύκλο ([4]). Κλείνοντας την αποτίμηση της σχετικής απόδοσης μεταξύ των υλοποιήσεων, εξετάζουμε 5 διαφορετικές περιπτώσεις υλοποίησης του απόλυτου της διαφοράς όπως αναφέρθηκαν στην ενότητα 6.1.2. Βάσει των σχετιζόμενων με τους πυρήνες *lseaiκ* και *cxHDκ* αποτελεσμάτων (κωδικοποιημένων ως “ενδογενείς” στην Εικ. 6.6), η λύση του `__usad` επιτυγχάνει την υψηλότερη απόδοση. Η καταμετρούμενη επιτάχυνση παραμένει σε σταθερά επίπεδα για όλους τους επεξεργαστές γραφικών και αναλυτικότερα, 1.0617x εν συγκρίσει με τη συνάρτηση *fabsf*, 1.1416x εν συγκρίσει με τον τριαδικό τελεστή (ABS_t), 1.0719x εν συγκρίσει με την υλοποίηση ακεραίων ολισθήσεων (ABS_s) και 1.1591x εν συγκρίσει με την ενδογενή εντολή `__vabsdiffu4` (υποστηριζόμενο στους επεξεργαστές αρχιτεκτονικής Kepler και Maxwell). Παρατηρούμε ότι παρόλο που η τελευταία SIMD εντολή έχει ρυθμοαπόδοση της τάξης των 160 λειτουργιών/κύκλο/πολυεπεξεργαστή στον *GTX 670*, when considering *lseaiκ* η λύση της συνάρτησης `__usad` αντιστοιχεί σε μόλις τέσσερις εντολές *sad.u32* οι οποίες έχουν την ίδια ρυθμοαπόδοση με τέσσερις εντολές *shl.b32* τις οποίες απαιτεί η SIMD εντολή. Ωστόσο, από τη μελέτη του PTX κώδικα προκύπτει ότι χρειάζονται 8 πρόσθετες εντολές *and.b32*, 2 πρόσθετες εντολές *shl.b32* και 6 πρόσθετες εντολές *or.b32* οι οποίες συνολικά δικαιολογούν το επίβαρο υπολογιστικό κόστος της SIMD εντολής.

Τέλος, για την αποτίμηση των απόλυτων τιμών επιδόσεων της προτεινόμενης

Εικόνα 6.7: Χρόνος εκτέλεσης *SIL-SEABI* σε ΜΕΓ.

υλοποίησης του αλγορίθμου *SIL-SEABI* καταμετρούμε το χρόνο εκτέλεσης των πυρήνων υπερ-δειγματοληψίας, παραλείποντας τις μεταφορές μνήμης προς και από τη διάταξη GPU (H2D / D2H). Για το σκοπό αυτό συγκρίνουμε την απόδοση με δικές μας υλοποιήσεις παρεμβολής ακεραίων 4 και 6 δειγμάτων οι οποίες συνιστούν δύο από τις απλούστερες δυνατές λύσεις υπερ-δειγματοληψίας. Συγκρίνουμε επίσης με την υλοποίηση σε επεξεργαστή γραφικών της μεθόδου *DPSR*, ως αντιπροσωπευτικής της αιχμής της τεχνολογίας. Τα αποτελέσματα καταγράφονται στην Εικ. 6.7 όπου σχηματίζεται ο μέσος χρόνος υπερ-δειγματοληψίας για κάθε ανάλυση από το σύνολο των δειγμάτων εικόνων XE (Ενότητα 7.1). Αξίζει να σημειωθεί ότι παρόλο που ο εξεργαστής *GTX 670* διαθέτει 24% περισσότερους επεξεργαστικούς πυρήνες από τον *GTX 960* επιτυγχάνει παρόμοιες επιδόσεις με τον τελευταίο, δηλαδή διαφορά μικρότερη του 10% ενώ είναι και 14.77% ταχύτερος όσον αφορά τον αλγόριθμο *SIL-SEABI*. Αυτό οφείλεται στον εσωτερικό σχεδιασμό της αρχιτεκτονικής Maxwell: κάθε Νηματικός Πολυεπεξεργαστής διαθέτει επεξεργαστικούς πυρήνες με πλήθος δύναμη του 2, λειτουργικές μονάδες με αποκλειστικό χρονοπρογραμματιστή, 8 μονάδες Φόρτωσης/Αποθήκευσης ανά 32 επεξεργαστικούς πυρήνες (αντί για 16 ανά 96 της αρχιτεκτονικής Kepler) και τέλος, μεγαλύτερο πλήθος καταχωρητών ανά επεξεργαστικό πυρήνα.

Όπως αναμενόταν, οι απλοί πυρήνες υπερ-δειγματοληψίας αποτελούν την ταχύτερη δυνατή λύση. Η δικυβική παρεμβολή για την υπερ-δειγματοληψία σε 3840×2160 χρειάζεται έως 0.4ms στον επεξεργαστή *GTX 670* και έως 1.14ms

στον *GTX 550 Ti* για την παραγωγή εικόνας ΥΕ μεγέθους 3840×2160 . Το φίλτρο 6 δειγμάτων είναι κατά 1.4 φορές βραδύτερο, το οποίο βρίσκεται σε συνέπεια με το λόγο των μεγεθών των πυρήνων φιλτραρίσματος. Συνεπώς, όταν η είσοδος ΧΕ έχει μέγεθος 1920×1080 δειγμάτων, οι απλοί πυρήνες υπερ-δειγματοληψίας επιτυγχάνουν απόδοση της τάξης των 698 πλαισίων ανά δευτερόλεπτο (π.α.δ., frames-per-second - fps) στον επεξεργαστή μέσης απόδοσης και 1718 π.α.δ. στους επεξεργαστές υψηλής απόδοσης. Επίσης αναμενόμενο είναι το γεγονός ότι οι μέθοδοι αυτές είναι κατά περίπου μία τάξη μεγέθους ταχύτερες από τον *SIL-SEABI* (5.1x και 7.6x όσον αφορά τα φίλτρα 6- και 4-δειγμάτων αντίστοιχα). Παρόλα αυτά, η προτεινόμενη υλοποίηση του αλγορίθμου *SIL-SEABI* παραμένει κατάλληλη για επεξεργασία εικόνων 1080p σε πραγματικό χρόνο καθώς επιτυγχάνει απόδοση της τάξης των 109 π.α.δ. στον επεξεργαστή *GTX 550 Ti* και 345 π.α.δ. στον *GTX 670*. Αντίθετα, η μέθοδος ΥΑ μέσω αποδιαμορφούμενων τμημάτων εικόνας Deformable Patches [101], απαιτεί τέσσερις τάξεις μεγέθους περισσότερο χρόνο κατά την υπερ-δειγματοληψία σε 176×144 και έως έξι τάξεις μεγέθους κατά την υπερ-δειγματοληψία σε 2160p. Το αποτύπωμα μνήμης της προτεινόμενης υλοποίησης του *SIL-SEABI* είναι επίσης σχετικά περιορισμένο και συγκρίσιμο με τα 46MB των απλών μεθόδων καθώς απαιτούνται μόλις 151MB μνήμης γραφικών και συνεπώς η προτεινόμενη υλοποίηση δύναται να εκτελεστεί σε όλους τους σύγχρονους επεξεργαστές γραφικών. Από την άλλη πλευρά, η μέθοδος *DPSR* είναι σημαντικά πιο περιοριστική καθώς απαιτεί 466MB μνήμης (176×144 δείγματα εξόδου) και έως 1452MB μνήμης για εισόδους μεγαλύτερης ευκρίνειας.

6.2 Επιτάχυνση του L-SEABI σε FPGA

Η προτεινόμενη επιτάχυνση σε υλικό περιλαμβάνει τη σχεδίαση αρχιτεκτονικής παράλληλης επεξεργασίας, την ανάπτυξη σε παραμετρικό κώδικα περιγραφής υλικού και την προετοιμασία για εφαρμογή του αλγορίθμου *L-SEABI* σε διάφορες μονάδες Συστοιχιών Πυλών Προγραμματιζόμενων στο Πεδίο (FPGAs). Στόχος της παραμετρικής αυτής υλοποίησης σε ποικίλλες μονάδες FPGA είναι η αποτίμηση της απόδοσης και της καταναλισκόμενης ισχύος του αλγορίθμου για διάφορες τιμές βαθμών παραλληλίας ανά FPGA. Η αρχιτεκτονική αυτή προέκυψε από τη συνεργασία με την ερευνητική ομάδα του εργαστηρίου Μικροϋπολογιστών και Ψηφιακών Συστημάτων του Εθνικού Μετσοβείου Πολυτεχνείου και στην παρούσα ενότητα παρέχεται μία σύντομη περίληψη της υλοποίησης και των αποτελεσμάτων αυτής.

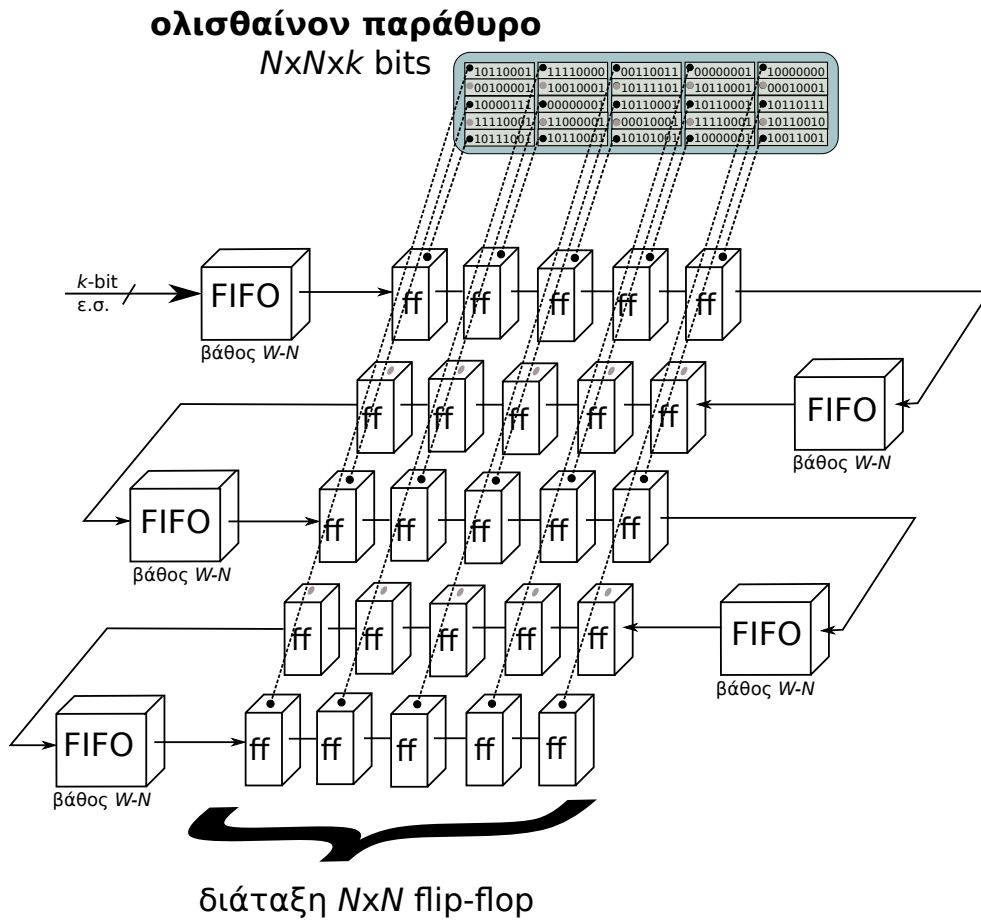
6.2.1 Σχεδίαση παράλληλης αρχιτεκτονικής για τον αλγόριθμο SIL-SEABI

Πέρα από τη βελτίωση της ποιότητας, ο προτεινόμενος αλγόριθμος ΥΑ ([31] και Ενότητα 5) στοχεύει στη δυνατότητα παραλληλοποίησης και το χαμηλό κό-

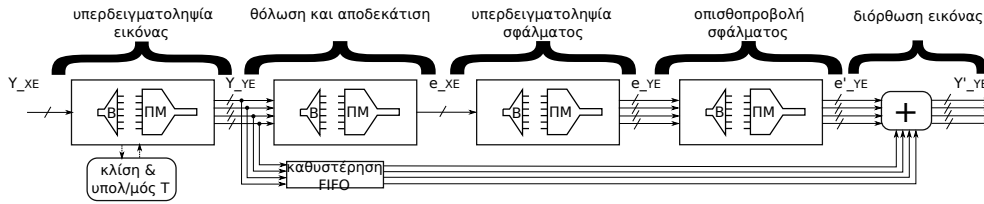
στος υλοποίησης σε υλικό, ειδικά σε επαναδιαρθρώσιμες πλατφόρμες με περιορισμένους πόρους μνήμης, όπως είναι τα FPGA. Συνεπώς, εν συγκρίσει με παρόμοιους και περισσότερο σύνθετους αλγορίθμους, ο *SIL-SEABI* αποφεύγει την περίπλοκη μαθηματική/στατιστική προσέγγιση, τις εξαρτήσεις μεταξύ των υπολογισμών, ή την αποθήκευση πρόσθετης πληροφορίας (π.χ. λεξικά). Αντί των παραπάνω, ο αλγόριθμος προωθεί την τοπικότητα και την κανονικότητα στους υπολογισμούς.

Η προτεινόμενη αρχιτεκτονική βασίζεται σε εκτεταμένη σωληναγωγή (pipelining) των υπολογισμών, τόσο σε επίπεδο εικονοστοιχείου, όσο και σε επίπεδο διεργασίας. Στηρίζεται στη συνεχή επεξεργασία της εικόνας από μία σωλήνωση μεγάλου βάθους η οποία περιλαμβάνει όλες τις απαιτούμενες αριθμητικές λειτουργίες για την ολοκλήρωση των διεργασιών του αλγορίθμου σε κάθε εικονοστοιχείο. Να σημειωθεί εδώ ότι με τον όρο *διεργασία*, αναφερόμαστε σε οποιαδήποτε στοιχειώδη και εννοιολογικά διακριτή λειτουργία μετασχηματισμού της εικόνας, όπως τη θόλωση, την αποδεκάτιση, τη δισδιάστατη συνέλιξη, τον υπολογισμό μερικών παραγώγων κ.λ.π. Εν γένει, το κόστος σε υλικό μιας τέτοιας σωληναγωγής σε επίπεδο εικονοστοιχείου περιορίζεται στο κόστος μίας μηχανής επεξεργασίας ενός εικονοστοιχείου και συνεπώς είναι ανεξάρτητο του μεγέθους της εικόνας. Σε ορισμένες περιπτώσεις, η σωλήνωση θα πρέπει επιπλέον να αποθηκεύσει ένα σύνολο εικονοστοιχείων (π.χ. για να υπολογίσει τις μερικές παραγώγους σε ένα τμήμα εικόνας), το πλήθος των οποίων ενδέχεται να είναι ανάλογο του πλάτους της εικόνας (π.χ. όταν η επεξεργασία γίνεται με σειρά σάρωσης πλεγματοειδούς εικόνας - raster scan order). Ο βέλτιστος σχεδιασμός της σωλήνωσης επιτρέπει τη διαρκή αξιοποίηση του συνόλου των πόρων. Στην περίπτωση αυτή, η ρυθμοαπόδοσή της είναι ένα εικονοστοιχείο ανά κύκλο ρολογιού και συνεπώς η επεξεργασία ολοκληρώνεται εντός $W \times H + L$ κύκλων, όπου W είναι το πλάτος, H το ύψος της εικόνας, και L ο λανθάνων χρόνος της σωλήνωσης. Με άλλα λόγια, η προτεινόμενη σωλήνωση σχεδιάστηκε ώστε να ολοκληρώνει μία επανάληψη/διεργασία του αλγορίθμου με κάθε ανάγνωση ριπής της εικόνας.

Για την επίτευξη της προαναφερθείσας ρυθμοαπόδοσης, παραλληλοποιούμε τους υπολογισμούς για κάθε εικονοστοιχείο, τόσο σε επίπεδο αριθμητικών πράξεων εντός των εξισώσεων, όσο και σε επίπεδο υπολογισμού μερικών γινομένων εντός μίας διεργασίας. Αναλυτικότερα, παραλληλοποιούμε τον υπολογισμό οριζόντιων και κατακόρυφων μερικών παραγώγων, τους υπολογισμούς των ενδιάμεσων εικονοστοιχείων κατά την προσαρμοστική υπερδειγματοληψία, τις διεργασίες διαχωρίσιμης συνέλιξης με χρήση του δισδιάστατου ισοδύναμου κ.λ.π. Η βασική ιδέα γύρω από την παραλληλοποίηση είναι η σύνδρομη επεξεργασία μιας τοπικής περιοχής της εικόνας και η ολοκλήρωση ενός μετασχηματισμού εικονοστοιχείου σχεδόν εξ ολοκλήρου σε έναν κύκλο. Για το σκοπό αυτό ενσωματώνουμε στη σωλήνωση μία μνήμη σειριακού σε παράλληλο, η οποία εισαγάγει ένα δείγμα ανά κύκλο ρολογιού (όσο και ο ρυθμός εισόδου της σωλήνωσης) και εξαγάγει $N \times N$ δείγματα ανά κύκλο. Η εικόνα 6.8 απεικονίζει την αρχιτεκτονική της προτεινόμενης ενδιάμεσης μνήμης η οποία συνδέει σειριακά N μονάδες μνήμης πρώτου δεδομένου μέσα-πρώτου δεδομένου έξω (first-in first-out - FIFO) με



Εικόνα 6.8: Προτεινόμενη αρχιτεκτονική διάταξη ανάγνωσης/επεξεργασίας σειριακού σε παράλληλο για τη διδιάστατη σάρωση πλεγματικής εικόνας με ολισθαίνον παράθυρο $N \times N$ εικονοστοιχείων.



Εικόνα 6.9: Η αρχιτεκτονική μίας μηχανής ΥΑ του αλγορίθμου *L-SEABI*. Διακρίνονται οι σωληνώσεις οι οποίες περιλαμβάνουν μνήμες μετατροπής σειριακού σε παράλληλο (B) και το επεξεργαστικό σκέλος των παράλληλων μονάδων (PUs). Ο ρυθμός επεξεργασίας ποικίλλει από ένα έως τέσσερα δείγματα ανά κύκλο αναλόγως της χωρικής ανάλυσης της εικόνας.

μία συστοιχία $N \times N$ καταχωρητών. Ορίζουμε το βάθος κάθε μνήμης FIFO σε $W - N$ και τροφοδοτούμε τη συνολική μονάδα με εικονοστοιχεία με σειρά σάρωσης πλεγματικής εικόνας. Η έξοδος $N \times N$ εικονοστοιχείων της ενδιάμεσης μνήμης επεξεργάζεται την εικόνα παρέχοντας ένα διακριτό “παράθυρο” $N \times N$ εικονοστοιχείων ανά κύκλο ρολογιού. Η προτεινόμενη σχεδίαση χρησιμοποιεί τον ελάχιστο αριθμό διασυνδέσεων (1 είσοδος και 1 έξοδος ανά υπομονάδα, με εξαίρεση την έξοδο όλης της μονάδας) και τον ελάχιστο αριθμό καταχωρητών (ίσο με τον αριθμό των παράλληλων θυρών εξόδου). Κατά συνέπεια, σε κάθε κύκλο ρολογιού προωθούνται στη σωλήνωση $N \times N$ εικονοστοιχεία, τα οποία επεξεργάζομαστε παράλληλα μέσω πολλαπλών αριθμητικών μονάδων διασυνδεδεμένων ανάλογα με την προς υλοποίηση εξίσωση. Συνολικά, η σωλήνωση μεγάλου βάθους αποτελούμενη από την ενδιάμεση μνήμη και τις αριθμητικές μονάδες εισαγάγει ένα εικονοστοιχείο ανά κύκλο και παράγει ένα εικονοστοιχείο ανά κύκλο με συνεχή ροή (εκτός από ορισμένα πολύ μικρά τακτά κενά διαστήματα στο τέλος κάθε οριζόντιας γραμμής της εικόνας). Ο κύριος λανθάνων χρόνος αφορά στο διάστημα που μεσολαβεί για τη συμπλήρωση των μνημών FIFO και σε σύγκριση με το συνολικό χρόνο επεξεργασίας είναι αμελητέος (0.4% κατά τη θόλωση μίας εικόνας 1920×1080 με ένα παράθυρο 5×5 εικονοστοιχείων).

Εκτός της σωληναγωγής σε επίπεδο εικονοστοιχείου, εφαρμόζουμε σωληναγωγή σε επίπεδο διεργασίας, διασυνδέοντας σειριακά τις 5 διακριτές μηχανές επεξεργασίας για τις ισάριθμες διεργασίες του αλγορίθμου *SIL-SEABI*. Η εικόνα 6.9 παρουσιάζει την επισκόπηση της αρχιτεκτονικής η οποία σχεδιάστηκε με προσανατολισμό “μετασχηματισμού συνεχούς ροής” (continuous flow transformation). Με άλλα λόγια, κάθε εικονοστοιχείο προωθείται άμεσα στο επόμενο στάδιο για περαιτέρω επεξεργασία ώστε να ολοκληρώνει τη συνολική επεξεργασία του από τον αλγόριθμο *SIL-SEABI* με ένα μόνο πέρασμα από τη σωλήνωση. Η προτεινόμενη αρχιτεκτονική επιτυγχάνει 100% αξιοποίηση πόρων επεξεργαζόμενη παράλληλα εκατοντάδες δείγματα εντός των παράλληλων μονάδων (PU) των πέντε διακριτών μηχανών επεξεργασίας. Κατά το πρώτο στάδιο (Εικ. 6.9), υπερδειγματοληπτούμε την εικόνα εισόδου βάσει της φάσης δόμησης του αλγορίθμου *L-SEABI*. Χρησιμοποιούμε μία μονάδα ενδιάμεσης μνήμης σειριακού σε παράλληλο (B) η οποία σε κάθε κύκλο ρολογιού προωθεί ένα πλα-

κίδιο εικονοστοιχείων ΧΕ διαστάσεων 4×4 σε δύο επεξεργαστικές μονάδες: α) στα υπερδειγματοληπτικά παράλληλα φίλτρα πεπερασμένης χρουστικής απόκρισης (Παράλληλες Μονάδες, ΠΜ) και β) σε μία σωλήνωση μικρότερου βάθους στην οποία συσσωρεύεται η τιμή των μερικών παραγώγων. Η υπομονάδα αυτή υπολογίζει την τιμή του κατωφλίου του αλγορίθμου βάσει της συνολικής διακύμανσης της εικόνας. Να σημειωθεί εδώ ότι χρησιμοποιούμε την τιμή της συνολικής διακύμανσης μίας λωρίδας για την πρόβλεψη της επόμενης, αποφεύγοντας έτσι μέρος των υπολογισμών. Εν συνεχεία, οι τιμές του κατωφλίου προωθούνται στην κύρια μονάδα η οποία εξαγάγει ένα πλακίδιο εικονοστοιχείων διαστάσεων 2×2 . Ο τετραπλασιασμός των μονοπατιών έχει το πλεονέκτημα της αποφυγής τομέων με διαφορετικές συχνότητες λειτουργίας καθώς και την αξιοποίηση του παραλληλισμού που προσφέρεται από το υλικό. Τα υπόλοιπα τέσσερα στάδια της σωλήνωσης αφορούν το στάδιο βελτίωσης του *SIL-SEABI*. Αναλυτικότερα, στο δεύτερο στάδιο πραγματοποιείται φιλτράρισμα και αποδεκάτιση της εικόνας ΥΕ με ρυθμό εξόδου ενός εικονοστοιχείου ανά κύκλο. Το σφάλμα των ανακατασκευασμένων εικονοστοιχείων προωθείται στο τρίτο στάδιο, όπου πραγματοποιείται προσαρμοστική υπερδειγματοληψία του σφάλματος παρόμοια με το πρώτο στάδιο (ωστόσο βάσει των κανόνων του *CxScale*). Τα αποτελέσματα προωθούνται στο τέταρτο στάδιο σε μορφή πλακιδίων 2×2 , όπου επεξεργαζόμαστε παράλληλα 4 περιοχές διαστάσεων 5×5 με ισάριθμους Λαπλασιανούς πίνακες έτσι ώστε να εξαγάγονται 4 εικονοστοιχεία ανά κύκλο. Κατά το τελικό στάδιο, τα 4 αυτά εικονοστοιχεία προστίθενται στα αντίστοιχα 4 της αρχικής ανακατασκευασμένης εικόνας ΥΕ, τα οποία αποθηκεύουμε σε μνήμη FIFO μετά το τέλος του πρώτου σταδίου. Καλούμε τον αναγνώστη να παρατηρήσει ότι η περιγραφείσα αρχιτεκτονική μπορεί να τροποποιηθεί ώστε να υποστηρίζει περισσότερες επαναλήψεις του αλγορίθμου *L-SEABI*. Μέσω μίας πρόσθετης μνήμης μπορούμε να αποθηκεύσουμε όλη την εικόνα εξόδου ${}^iY_{HR}$ από τη σωλήνωση και για κάθε επανάληψη i , τροφοδοτώντας τη κατευθείαν στο δεύτερο στάδιο.

6.2.2 Εξερεύνηση χώρου σχεδίασης και αποτελέσματα υλοποίησης

Η προτεινόμενη αρχιτεκτονική αναπτύχθηκε σε γλώσσα VHDL με παραμετρική σχεδίαση ώστε να επιτρέπει την προσαρμογή στις απαιτήσεις της εκάστοτε εφαρμογής αλλά και τη δοκιμή του αλγορίθμου *L-SEABI* σε ποικίλλες μονάδες FPGA. Ανάμεσα στις επιλέξιμες παραμέτρους είναι το εύρος των διαύλων δεδομένων, το μέγεθος της εικόνας, οι διαστάσεις των πυρήνων συνέλιξης καθώς και ο βαθμός παραλληλίας του αλγορίθμου στο επίπεδο μίας εικόνας. Ο παραλληλισμός σε επίπεδο εικόνας αποτελεί ένα ακόμη χαρακτηριστικό της σχεδίασης σε υλικό, υπό την έννοια του αριθμού των εικόνων ή των λωρίδων μίας εικόνας στα οποία διενεργείται επεξεργασία παράλληλα. Εν προκειμένω, χρησιμοποιούμε P παράλληλες σωληνώσεις, τροφοδοτώντας την κάθε μία από αυτές με μία οριζόντια λωρίδα εικόνας κατακόρυφης διάστασης $H/P + O$ εικονοστοιχείων όπου H η κατακόρυφη διάσταση της αρχικής εικόνας και O το πλήθος των γραμμών επικάλυψης μεταξύ διαδοχικών λωρίδων. Να σημειώσουμε εδώ ότι για να είναι ο

Πίνακας 6.2: Κόστος πόρων διάταξης FPGA έναντι χρόνου εκτέλεσης (στις διατάξεις εισαγάζεται μία εικόνα διαστάσεων 1920×1080 και επιλέγονται διάφοροι βαθμοί παραλληλίας για παράγοντα υπερδειγματοληψίας $f = 2$)

συστοιχία	P	LUT	RAMB18	κατανάλωση (W)	χρόνος (ms)
xc5vlx30t	1	3753	64	0.9	8.3
xc6vlx75t	1	3493	64	2.3	6.8
xc7a100t	1	3501	64	0.5	8.2
xcku035	1	3637	64	1	5
xc5vlx110t	4	22%	89%	3	2.7
xc5vlx330t	8	15%	79%	7.5	1.5
xc6vlx75t	3	22%	61%	3	3
xc6vlx240t	11	26%	85%	7.8	1
xc6vlx550t	18	18%	91%	11	.71
xc7a100t	3	16%	71%	1	3.5
xc7vx485t	16	19%	49%	5.5	.70
xc7v2000t	29	9%	71%	11	.48
xcku035	12	24%	68%	6.2	.65

παράλληλος αλγόριθμος λειτουργικά ισοδύναμος με τον αρχικό, θα πρέπει να επεκταθούν τα όρια της κάθε λωρίδας έτσι ώστε η συνέλιξη να εκτελείται αδιάλειπτα. Στην παρούσα διατριβή, θεωρούμε ότι κάθε εικονοστοιχείο χρειάζεται 8-bit αποθηκευτικό χώρο, η εικόνα εισόδου έχει διαστάσεις 1920×1080 εικονοστοιχεία και $O = 9$ (εξαιτίας των διαστάσεων στους πυρήνες συνέλιξης του $L - SEABI$). Επίσης, δε λαμβάνουμε υπόψιν τις απαιτήσεις σε ακροδέκτες εισόδου/εξόδου για να εστιάσουμε στο κόστος του ίδιου του αλγορίθμου (θεωρούμε ότι την εισοδο/έξοδο διαχειρίζονται εξωτερικές μονάδες, εξαρτώμενες από την εκάστοτε εφαρμογή). Η προτεινόμενη αρχιτεκτονική υλοποιήθηκε σε 10 διατάξεις FPGA, με τις οποίες εξετάζονται ουσιαστικά τέσσερις γενιές τεχνολογικής εξέλιξης (πρόκειται για διατάξεις Xilinx Virtex 5, Virtex 6, Artix και Virtex 7, καθώς και Ultrascale). Για την εξερεύνηση των επιδόσεων του $L-SEABI$, μεταβάλλουμε το βαθμό παραλληλίας P από 1 έως 29, ανάλογα με τις δυνατότητες της κάθε διάταξης.

Τα αποτελέσματα της υλοποίησης παρουσιάζονται συνοπτικά στον Πίνακα 6.2, ο οποίος περιλαμβάνει τους απαιτούμενους πόρους σε πίνακες αναζήτησης (Look-Up Tables - LUTs) και μονάδες μνήμης RAMB18, καθώς και το συνολικό χρόνο εκτέλεσης μαζί με την κατανάλωση στη μέγιστη επιτευχθείσα συχνότητα για την υπερδειγματοληψία μίας εικόνας διαστάσεων 1920×1080 σε 3840×2160 . Στο άνω τμήμα του πίνακα 6.2, αναγράφονται οι απόλυτοι δείκτες επιδόσεων για μία σωλήνωση του αλγορίθμου $L-SEABI$ όταν αυτή υλοποιείται σε χαμηλής αποδοτικότητας και κατανάλωσης συστοιχίες οι οποίες αντιστοιχούν σε 4 γενιές εξέλιξης (Virtex 5, 6, 7, και Ultrascale). Οι υλοποιήσεις αυτές επιτυγχάνουν απόδοση της τάξης των 120-200 π.α.δ. καταναλώνοντας μόλις 0.5-2.3 W (εκτίμηση

μέσω Xilinx Power Analyzer). Η στατική κατανάλωση κάθε συστοιχίας εξαρτάται από τη λιθογραφική κλίμακα και το μέγεθος του πλακιδίου των κυκλωμάτων, ενώ η δυναμική κατανάλωση καθορίζεται από τη συχνότητα λειτουργίας και το ποσοστό αξιοποίησης των συνολικών πόρων. Η μέγιστη συχνότητα λειτουργίας για μία σωλήνωση κυμαίνεται από 250-420 MHz αναλόγως της συστοιχίας και ελαττώνεται στα 185-300 MHz όταν τοποθετούνται πολλαπλές μηχανές σωλήνωσης στην ίδια συστοιχία (ωστόσο οι συχνότερες τιμές είναι τα 200-220 MHz). Στο κάτω τμήμα του πίνακα 6.2 παρουσιάζονται τα αποτελέσματα της τοποθέτησης πολλαπλών σωληνώσεων σε κάθε συστοιχία. Η δεύτερη στήλη αναφέρει το μέγιστο πλήθος ανά συστοιχία, οι στήλες 3 και 4 την αξιοποίηση των πινάκων αναζήτησης (LUTs) και των μνημών (RAMBs). Να σημειωθεί εδώ ότι ο περιοριστικός παράγοντας σε συστοιχίες μικρής και μεσαίας ολοκλήρωσης είναι η ενσωματωμένη μνήμη, ενώ σε συστοιχίες υψηλής ολοκλήρωσης οι ακροδέκτες εισόδου/εξόδου.

Βάσει των ανωτέρω αποτελεσμάτων, η αύξηση του βαθμού παραλληλίας P σε κάθε γενιά FPGA οδηγεί σε σχεδόν γραμμική επιτάχυνση του αλγορίθμου. Η απόδοση περιορίζεται αφενός από την ελάττωση της μέγιστης συχνότητας (εξαιτίας του υψηλού ποσοστού αξιοποίησης πόρων) και το επίβαρο κόστος των επικαλυπτόμενων τμημάτων εικόνας σε διαδοχικές λωρίδες το οποίο απαιτείται προκειμένου να υπάρξει λειτουργικά ισοδύναμο αποτέλεσμα. Στη συστοιχία Virtex 7 2000T (τη δαπανηρότερη όσων εξετάστηκαν), η αρχιτεκτονική επιτυγχάνει 2083 π.α.δ. για $P = 29$. Σε αυτή την περίπτωση, το επίβαρο κόστος αυξάνει σε περίπου 25% εξαιτίας του εκτεταμένου βαθμού επικάλυψης μεταξύ των λωρίδων. Σε υλοποιήσεις πολλαπλών σωληνώσεων, η κατανάλωση κυμαίνεται από 1 έως 11 W, ενώ η ενέργεια ανά εικόνα περίπου υποδιπλασιάζεται στις συστοιχίες νεότερης γενιάς (π.χ. Virtex 7 έναντι Virtex 5/6). Χρησιμοποιώντας μία μέσου κόστους συστοιχία όπως η xc6vlx240t, αρκεί για να επιτύχουμε επιδόσεις της τάξης των 1000 π.α.δ. με λιγότερα από 8 W κατανάλωσης, ενώ η χαμηλού κόστους xc7a100t αποδίδει 285 π.α.δ. με μόλις 1 W. Συνεπώς η προτεινόμενη μηχανή επιτάχυνσης του αλγορίθμου μπορεί να υποστηρίξει πληθώρα εφαρμογών και αναλύσεων, επιτυγχάνοντας επεξεργασία σε πραγματικό χρόνο.

Σε σύγκριση με άλλες υλοποιήσεις αλγορίθμων ΥΑ σε FPGA από τη βιβλιογραφία, η προτεινόμενη υλοποίηση μίας σωλήνωσης στη διάταξη xc6vlx240t απαιτεί περίπου $5 \times$ λιγότερους πόρους από τη μονάδα στη [74]. Η υλοποίηση τριών σωληνώσεων έχει παραπλήσιο κόστος σε πόρους με τη μονάδα στη [69], ωστόσο επιτυγχάνει μία τάξη μεγέθους ταχύτερη εκτέλεση στην ίδια διάταξη (xc6vlx75t). Η υλοποίηση 11 παράλληλων σωληνώσεων στη διάταξη xc6vlx240t απαιτεί παρόμοιο πλήθος πινάκων αναζήτησης με την τετραπύρρηνη υλοποίηση στη [65], ολοκληρώνοντας την υπερδειγματοληψία σε μία τάξη μεγέθους λιγότερο χρόνο από την τελευταία (διάταξη FPGA Altera Aria II). Τέλος, εν συγκρίσει με πιο απαιτητικές υλοποιήσεις ΥΑ, η προτεινόμενη αρχιτεκτονική είναι σημαντικά ταχύτερη και με πολύ μικρότερο κόστος σε πόρους. Για παράδειγμα, για την υπερδειγματοληψία εικόνων σε ανάλυση QCIF μία μηχανή σωλήνωσης στη συστοιχία Virtex 5 αποδίδει πάνω από $400 \times$ περισσότερα π.α.δ. με περίπου $3 \times$ λιγότερους πίνακες

αναζήτησης σε σχέση με την ΥΑ με εκτίμηση κίνησης στη [85]. Ωστόσο, πρέπει να σημειωθεί εδώ ότι γενικά η ΥΑ με εκτίμηση κίνησης αποδίδει αποτελέσματα υψηλότερης ποιότητας.

6.3 Σύγκριση υλοποιήσεων σε CPU, GPU και FPGA

Σε αυτή την ενότητα, θα αποτιμηθεί συγκριτικά η απόδοση των προτεινόμενων υλοποιήσεων επιτάχυνσης του αλγορίθμου, λαμβάνοντας υπόψιν την ποιότητα των παραγόμενων αποτελεσμάτων, την καταναλισκόμενη ισχύ και το κόστος της πλατφόρμας υλοποίησης. Για το σκοπό αυτό, θα συγκρίνουμε την απόδοση των υλοποιήσεων στους προαναφερθέντες επεξεργαστές γραφικών και σε διάφορα ολοκληρωμένα FPGA έχοντας ένα κοινό σημείο αναφοράς: την καταμετρούμενη απόδοση σε πολυπύρηνους επεξεργαστές γενικού σκοπού αρχιτεκτονικής x86_64. Καθόλη τη διάρκεια των πειραμάτων, θεωρούμε ότι κάθε εικόνα ΧΕ υπερ-δειγματοληπτείται κατά παράγοντα $f=2$ και στις δύο διευθύνσεις με σκοπό να παραχθούν οι εικόνες ΥΕ, οι οποίες έχουν το ίδιο μέγεθος με την πηγαία εικόνα.

6.3.1 L-SEABI: αποτελέσματα ποιότητας

Οι υλοποιήσεις επιτάχυνσης του αλγορίθμου *SIL-SEABI* όπως αυτές παρουσιάστηκαν στις ενότητες 6.1.2, και 6.2, περιλαμβάνουν τόσο αριθμητικές (π.χ. μέσω των σταθερών της Γκαουσιανής συνάρτησης φιλτραρίσματος) όσο και αλγοριθμικές προσεγγίσεις, όπως τον αδρομερή υπολογισμό των δειγμάτων ΥΕ σε ΗΔ θέσεις (υλοποίηση GPU), ή τον υπολογισμό της συνολικής μεταβλητότητας χωρίζοντας την εικόνα σε οριζόντιες λωρίδες (υλοποίηση FPGA). Για την αποτίμηση του αντικτύπου αυτών των προσεγγίσεων, συγκρίνουμε την ποιότητα του αποτελέσματος των επιταχυμένων υλοποιήσεων με το αποτέλεσμα αναφοράς της υλοποίησης σε επεξεργαστή γενικού σκοπού (ακριβώς όπως στον Πίνακα 7.1). Βάσει των αποτελεσμάτων (Πίνακας 6.3) παρατηρούμε ότι η ποιότητα - όπως αυτή μετράται από τις αντικειμενικές μετρικές MSSIM και BRISQUE - διατηρείται στα επίπεδα της υλοποίησης αναφοράς τόσο για τις υλοποιήσεις σε GPU, όσο και για τις υλοποιήσεις σε FPGA. Όπως αναμενόταν, η αύξηση του πλήθους των λωρίδων στην υλοποίηση σε FPGA προκαλεί μία μικρή ελάττωση ποιότητας, η οποία ωστόσο παραμένει αμελητέα ακόμη και στην περίπτωση της επεξεργασίας εικόνων QCIF με 29 οριζόντιες λωρίδες.

6.3.2 Αποτελέσματα επιτάχυνσης L-SEABI: ανάλυση υλοποίησης CPU

Για τις ανάγκες της αποτίμησης της εκτέλεσης του αλγορίθμου σε πολυπύρηνους επεξεργαστές γενικού σκοπού, χρησιμοποιούμε ένα πλήρως ανυσμα-

Πίνακας 6.3: Επιδόσεις Ποιότητας των υλοποιήσεων του αλγορίθμου *SIL-SEABI* σε CPU, GPU και FPGA.

Ποιότητα Υλοποιήσεων <i>SIL-SEABI</i>						
Πλατφόρμα:	CPU	GPU	FPGA			
Μέγεθος ΥΕ	(Ref.)	(Opt.)	8 λρδ.	16 λρδ.	29 λρδ.	
MSSIM	QCIF	0.9086	0.9032	0.9085	0.9084	0.9074
	CIF	0.8743	0.8701	0.8740	0.8744	0.8740
	SD1	0.9679	0.9681	0.9703	0.9702	0.9704
	720p	0.9567	0.9568	0.9593	0.9592	0.9592
	1080p	0.9784	0.9784	0.9798	0.9797	0.9797
	2160p	0.9879	0.9871	0.9879	0.9879	0.9879
μ.τ.	0.9456	0.9440	0.9467	0.9467	0.9464	
BRISQUE	QCIF	43.5880	41.5805	42.5290	40.9607	40.4722
	CIF	36.1080	35.9498	36.4587	36.4763	36.8599
	SD1	44.4549	41.8728	44.0211	44.2512	44.0111
	720p	43.5580	41.7760	43.3747	43.4144	43.4126
	1080p	42.0595	41.6913	42.0178	42.0280	42.0207
	2160p	57.9444	55.9248	56.8053	56.9504	57.0885
μ.τ.	44.1252	43.1269	44.2011	44.0135	43.9775	

τοποημένο μοντέλο του *SIL-SEABI* το οποίο αξιοποιεί τον ενδογενή πολυνηματισμό των συναρτήσεων του υπολογιστικού πακέτου MATLAB. Επιπρόσθετα, εξετάζουμε και την υλοποίηση μονού προγράμματος-πολλαπλών δεδομένων (MPPD, Single-program multiple data - *spm*) στο οποίο ο αριθμός των νημάτων-εργατών είναι ίσος με τον αριθμό των διαθέσιμων πυρήνων του επεξεργαστή. Σε αυτή την υλοποίηση η εικόνα χωρίζεται σε κατακόρυφα τμήματα πλήθους ίσου με τον αριθμό των εργατών και η ανταλλαγή μηνυμάτων κατά τη διάρκεια της επεξεργασίας πραγματοποιείται με τις συναρτήσεις *labSend* και *labReceive* του MATLAB. Κατ' αυτόν τον τρόπο, ο αλγόριθμος *SIL-SEABI* υλοποιείται στο διπύρηνο επεξεργαστή *Atom 330 dual-core*, τον τετραπύρηνο επεξεργαστή *Core i5-3470* και τον οκταπύρηνο επεξεργαστή *FX-8120*. Να σημειωθεί ότι στις περιπτώσεις εκτέλεσης *spm* κώδικα, ο χρόνος υπολογισμού καταμετράται έως και το βήμα επανασύστασης της εικόνας από τις κατακόρυφες λωρίδες.

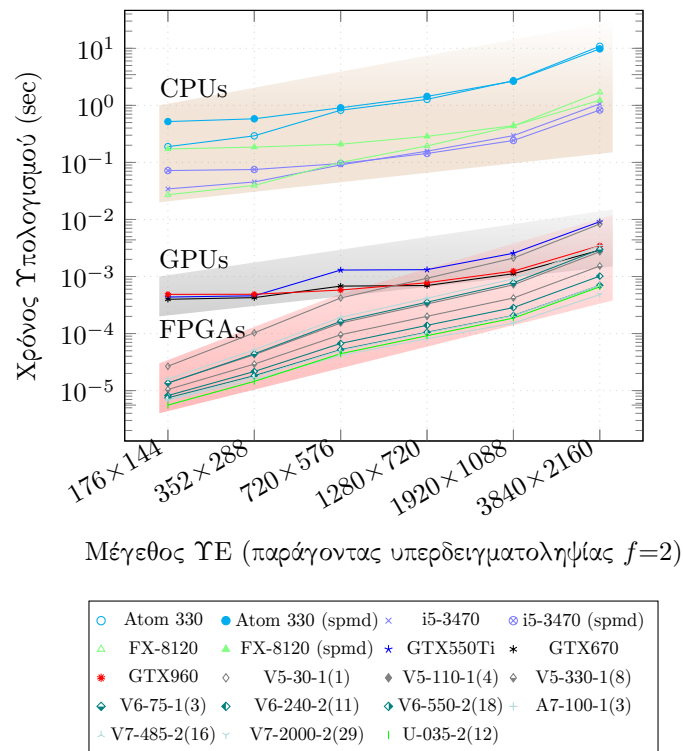
Τα αποτελέσματα των χρόνων υπολογισμού απεικονίζονται γραφικά για διάφορες τιμές ευκρίνειας εξόδου στην Εικ. 6.10 για όλες τις περιπτώσεις CPU, GPU και FPGA που εξετάστηκαν. Ο πίνακας 6.4 παρουσιάζει και την επιτευχθείσα επιτάχυνση ανά πλατφόρμα σε σχέση με την ταχύτερη υλοποίηση σε CPU προκειμένου ο αναγνώστης να αποκτήσει μία συνοπτική αλλά ακριβέστερη εικόνα των αποτελεσμάτων.

Όπως ήταν αναμενόμενο, ο επεξεργαστής *Atom 330* με μόλις 1 MB λανθάνουσας μνήμης L2 και έλλειψη υποστήριξης εκτέλεσης εντολών εκτός σειράς ή κερδοσκοπικής εκτέλεσης (*speculative execution*) επιτυγχάνει τις χαμηλότερες επιδόσεις - περίπου 10 sec για την υπερ-δευματοληψία σε ευκρίνεια UHD. Από

Πίνακας 6.4: Επιδόσεις Επιτάχυνσης του αλγορίθμου *SIL-SEABI* σε CPU, GPU και FPGA.

Επιτάχυνση Υλοποιήσεων SIL-SEABI							
Πλατφόρμα	Μέγεθος Εικόνας ΥΕ ($f = 2$)						
	QCIF	CIF	SD1	720p	1080p	2160p	μ.τ.
i5-3470	1	1	1	1	1	1	1
FX-8120	1.26	1.15	0.90	0.73	0.55	0.49	0.85
GTX550Ti	78.54	98.95	69.70	108.79	94.68	90.66	90.22
GTX670	86.04	106.5	133.2	206.3	216.4	285.8	172.4
GTX960	70.58	93.01	156.1	184.6	194.9	238.2	156.2
V5-30-1(1)	1283	440.8	215.1	153.9	115.1	99.5	384.6
V5-110-1(4)	2517	1065	591.2	435.2	338.6	304.8	875.3
V5-330-1(8)	3281	1558	948.7	714.9	576.2	542.7	1270
V6-75-1(3)	2530	1027	553.3	404.5	311.5	277.5	850.6
V6-240-2(11)	4191	2101	1347	1029	848.1	814.6	1722
V6-550-2(18)	4589	2485	1733	1358	1167	1181	2086
A7-100-1(3)	2161	877.0	472.6	345.5	266.1	237.0	726.5
V7-485-2(16)	4905	2610	1783	1389	1181	1179	2174
V7-2K-2(29)	5039	2905	2191	1764	1588	1717	2534
U-035-2(12)	6160	3132	2037	1564	1297	1256	2574

την άλλη πλευρά, ο τετραπύρηνος *i5-3470* χρειάζεται κάτω από 1 δευτερόλεπτο σε όλες σχεδόν τις περιπτώσεις και γι' αυτό θα αποτελέσει το μέτρο σύγκρισης των επιδόσεων επιτάχυνσης. Κατά την υπερ-δειγματοληψία σε ευκρίνεια SD1 αλλά και υψηλότερα μεγέθη, ο *i5-3470* παράγει την εικόνα ΥΕ ταχύτερα από τον *FX-8120* εξαιτίας των καθυστερήσεων στο υποσύστημα μνήμης του τελευταίου. Σε ένα διαφορετικό επίπεδο συγκρίσεων, παρατηρούμε ότι η υλοποίηση σε κώδικα *spmd* αυξάνει τις επιδόσεις του ενδογενούς μοντέλου πολυνηματισμού του MATLAB κατά 1.09x και έως 1.28x κατά την υπερ-δειγματοληψία σε 720p ή υψηλότερες αναλύσεις στον επεξεργαστή *i5-3470* (και έως 1.38x για υπερ-δειγματοληψία 2160p στους άλλους επεξεργαστές). Σε χαμηλότερες αναλύσεις, το επίβαρο κόστος επικοινωνίας είναι πολύ υψηλότερο του χρόνου που χρειάζεται ουσιαστικά για τους αριθμητικούς υπολογισμούς. Επομένως ο ενδογενής πολυνηματισμός των συναρτήσεων του MATLAB είναι στις περιπτώσεις αυτές προτιμότερος. Κατά την επεξεργασία εικόνων ΧΕ πολύ μεγαλύτερου μεγέθους ήτοι 6750×6750 δειγμάτων, ο χρόνος υπολογισμού αυξάνει σχεδόν γραμμικά - για παράδειγμα το *spmd* μοντέλο στον επεξεργαστή *FX-8120* ολοκληρώνει τους υπολογισμούς σε 21.39 sec.



Εικόνα 6.10: Συνοπτική εικόνα της απόδοσης του αλγορίθμου *SIL-SEABI* ανά πλατφόρμα επιτάχυνσης και ανάλυση αποτελέσματος: Απαιτούμενος χρόνος υπερδειγματοληψίας εικόνας ΧΕ κατά παράγοντα $f = 2$.

6.3.3 Αποτελέσματα επιτάχυνσης L-SEABI: ανάλυση υλοποίησης σε ΜΕΓ

Συγκρίνοντας τα αποτελέσματα της επιτάχυνσης σε GPU με αυτά της ταχύτερης υλοποίησης σε επεξεργαστή γενικού σκοπού (δηλαδή αυτά του επεξεργαστή *i5-3470* στον Πίνακα 6.4), καταμετρούμε ελάχιστη επιτάχυνση 69.7x στον μέσης αποδοτικότητας *GTX 550Ti* για εικόνες ΥΕ μεγέθους 720×576 δειγμάτων. Ο υψηλής αποδοτικότητας *GTX 670* επιτυγχάνει επιτάχυνση έως και κατά 285.8x καθώς κατά μέσο όρο χρειάζεται 2.89 ms προκειμένου να υπερ-δειγματοληπτήσει σε μέγεθος 2160p. Είναι προφανές ότι περαιτέρω αύξηση του μεγέθους των εικόνων ΧΕ περιορίζει το εύρος των καρτών γραφικών που είναι κατάλληλες για τόσο μεγάλο επεξεργαστικό φορτίο και αυτό συμβαίνει εξαιτίας περιορισμών στη συνολικά διαθέσιμη μνήμη. Επί παραδείγματι, ο επεξεργαστής *GTX 960* - η πλακέτα ενσωμάτωσης του οποίου διαθέτει 4GB μνήμης - χρειάζεται 347.55 ms για την υπερ-δειγματοληψία εικόνων ΧΕ μεγέθους 6750×6750 δειγμάτων κατά παράγοντα 2.

6.3.4 Αποτελέσματα επιτάχυνσης L-SEABI: ανάλυση υλοποίησης σε FPGA

Για λόγους διευκόλυνσης της συζήτησης γύρω από τα αποτελέσματα του πίνακα 6.4 και της Εικ. 6.10, ακολουθούμε τον εξής συμβολισμό: μία μηχανή στο *xc5vlx30t-1* FPGA θα αναγράφεται ως V5-30-1 (1), 11 μηχανές στο *xc6vlx240t-2* ως V6-240-2 (11), 16 μηχανές στο *xc7vx485t-2* ως V7-485-2 (16), 12 μηχανές στο *xc6ku035-2* ως U-035-2 (12) και ούτω καθεξής. Όπως φαίνεται στον πίνακα 6.4 και την Εικ. 6.10, η επιτευχθείσα επιτάχυνση σε οποιοδήποτε ολοκληρωμένο FPGA είναι πολύ υψηλότερη από όλες τις προαναφερθείσες υλοποιήσεις, ειδικά στην περίπτωση εικόνων ΧΕ 88×72 δειγμάτων στις οποίες ο λόγος των λειτουργιών μνήμης προς αυτόν των αριθμητικών υπολογισμών είναι υψηλός: ακόμη και μία μηχανή στο ολοκληρωμένο *xc5vlx30t-1* είναι σε θέση να υπερ-δειγματοληπτήσει εικόνα 88×72 δειγμάτων σε 176×144 εντός μόλις 0.026 ms, δηλαδή επιτυγχάνει επιτάχυνση κατά 1283x (Πίνακας 6.4).

Κατά την υπερ-δειγματοληψία σε 1080p ή μεγαλύτερα μεγέθη ΥΕ, η υλοποίηση μίας μηχανής στο χαμηλής αποδοτικότητας ολοκληρωμένο της γενιάς Virtex 5 επιτυγχάνει επιδόσεις παραπλήσιες αυτών του *GTX 550Ti*, όπως συμβαίνει και με τις υλοποιήσεις 3 μηχανών στα ολοκληρωμένα *xc7a100t-1*, *xc6vlx75t-1* και την υλοποίηση 4 μηχανών στο *xc5vlx110t-1* εν συγκρίσει με αυτές των γραφικών επεξεργαστών υψηλής αποδοτικότητας κατά την υπερ-δειγματοληψία σε μέγεθος 2160p. Η υλοποίηση 12 παράλληλων μηχανών στο ολοκληρωμένο *xc6ku035-2* παρέχει τις υψηλότερες επιδόσεις για εικόνες ΥΕ μεγέθους 176×144 δειγμάτων ενώ η απόδοση της υλοποίησης των 29 μηχανών στο *xc7v2000t-2* κυριαρχεί κατά την υπερ-δειγματοληψία σε μέγεθος 2160p από 1080p (επιταχύνσεις κατά 6160x και 1716x αντιστοίχως βάσει του Πίνακα 6.4). Να σημειωθεί εδώ ότι περαιτέρω αύξηση του μεγέθους των εικόνων ΧΕ σε 6750×6750 δείγματα περιορίζει το εύρος

των κατάλληλων FPGA ολοκληρωμένων αλλά και το βαθμό παραλληλίας, καθώς μία και μόνο μηχανή απαιτεί σε αυτή την περίπτωση 252 RAMBs μονάδες μνήμης. Συν τοις άλλους, το διάστημα μεταξύ των επιδόσεων GPU και FPGA περιορίζεται - παρόλο που τα τελευταία διατηρούν το πλεονέκτημά τους. Για παράδειγμα το xc6vlx240t-2, το οποίο σε αυτές τις περιπτώσεις υποστηρίζει 3 παράλληλες μηχανές, παράγει την εικόνα ΥΕ μεγέθους 13500×13500 σε έως και 80.35 ms. Σε ολοκληρωμένα υψηλότερης αποδοτικότητας είναι εφικτές ακόμη και επιδόσεις πραγματικού χρόνου όπως στο xc7vx485t-2, το οποίο με 8 παράλληλες μηχανές επιτυγχάνει επιδόσεις 33 π.α.δ. (χρόνος 28.9 ms με χρήση 2016 RAMB μονάδων μνήμης). Αξίζει να σημειωθεί ότι επιδόσεις αυτής της τάξης αντιστοιχούν σε επεξεργασία πραγματικού χρόνου εικόνων μεγέθους 182 Mpixels, δηλαδή μεγέθους πάνω από 2 φορές αυτού των επικείμενων οθονών τεχνολογίας 11K (ανάλυσης 11520×6480).

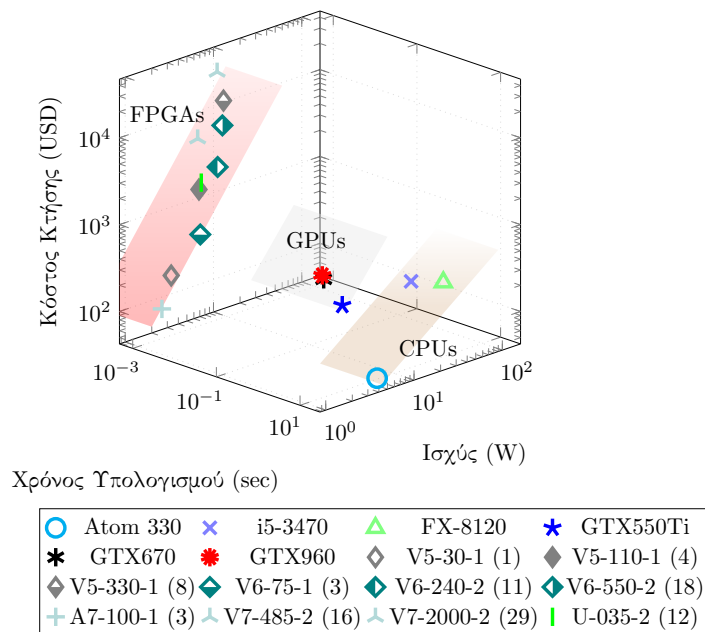
Αποτιμώντας συνολικά την προαναφερθείσα πειραματική ανάλυση, οι επεξεργαστές γενικού σκοπού επιτυγχάνουν επιδόσεις υπερ-δειγματοληψίας οι οποίες κατ' απόλυτη τιμή κυμαίνονται μεταξύ $2.72 \mu\text{s}/\text{δείγμα}$ στον διπύρηνο *Atom 330* έως $0.4 \mu\text{s}/\text{δείγμα}$ στους επεξεργαστές υψηλότερης αποδοτικότητας. Οι επεξεργαστές γραφικών παρέχουν κατά δύο τάξεις μεγέθους υψηλότερες επιδόσεις (0.0048 και έως $0.0038 \mu\text{s}/\text{δείγμα}$) ενώ στις συστοιχίες FPGA, η παραμετροποιήσιμη σχεδίαση της αρχιτεκτονικής σε συνδυασμό με τον υψηλό βαθμό παραλληλίας δύνανται να εκτοξεύσουν τις επιδόσεις σε επίπεδα υψηλότερα και από αυτά των GPU (κατά μία τάξη μεγέθους, ήτοι $0.12 \text{ ns}/\text{δείγμα}$ στο ολοκληρωμένο 7-2000-2).

6.3.5 Σύγκρισεις πλατφόρμας υλοποίησης: ισχύς και απόδοση

Έχοντας ως στόχο να ενισχύσουμε τη διεξοδικότητα της πειραματικής ανάλυσης, αποτιμούμε επιπλέον τις επιδόσεις από κοινού με την καταναλισκόμενη ισχύ, για την εκτέλεση του αλγορίθμου *SIL-SEABI* σε κάθε πλατφόρμα υλοποίησης. Όσον αφορά τους επεξεργαστές γενικού σκοπού και τους επεξεργαστές γραφικών, διενεργήθηκε μέτρηση της καταναλισκόμενης ισχύος σε επαναληπτική εκτέλεση του αλγορίθμου (μέγιστος αριθμός 10^6 επαναλήψεων, έως ότου η διακύμανση κατέστη αμελητέα) με τη βοήθεια ενός μετρητή ισχύος σε θερμοκρασία περιβάλλοντος 25 βαθμών Κελσίου και ελήφθη ο μέσος όρος των αποτελεσμάτων για όλα τα μεγέθη εικόνων. Όσον αφορά τις συστοιχίες FPGA, η καταναλωση ισχύος εκτιμήθηκε μέσω του εργαλείου *Xilinx Power Estimator (XPE)* στην ίδια θερμοκρασία περιβάλλοντος. Τα αποτελέσματα παρουσιάζονται γραφικά στην Εικ. 6.11 σε αντιπαραβολή με τον απαιτούμενο χρόνο υπερ-δειγματοληψίας από 1080p σε 2160p. Το γράφημα 6.11 απεικονίζει επίσης συγκεντρωτικά κατά προσεγγιστικό τρόπο τις καταμετρούμενες επιδόσεις έτσι ώστε να καθίσταται εμφανέστερη η διαφορά των επιδόσεων μεταξύ υλοποίησης σε CPU, GPU και FPGA μέσω σχιασμένων περιοχών (καφέ περιοχή για CPU, γκρι περιοχή για GPUs και ανοιχτόχρωμη κόκκινη για συστοιχίες FPGA).

Ανάμεσα στους εξετασθέντες επεξεργαστές, ο *Atom 330* εμφανίζει κατανά-

Αλγόριθμος *SIL-SEABI* - Χρόνος, Καταναλισχόμενη Ισχύς και Κόστος Κτήσης (1080p → 2160p)



Εικόνα 6.11: Χρόνος και Ισχύς κατά τη διεργασία ΥΑ μέσω του αλγορίθμου *SIL-SEABI*. Στο γράφημα περιλαμβάνονται διάφορα ολοκληρωμένα κυκλώματα μαζί με το αντίστοιχο κόστος κτήσης τους (σε USD) (ο συμβολισμός εξηγείται στο κυρίως κείμενο).

λωση παρόμοια με αυτή των ολοκληρωμένων FPGA: καταναλώνει περίπου 3 W σε κατάσταση ηρεμίας και συνολικά μεταξύ 5.1 W (υπερ-δειγματοληψία QCIF) έως 6.5 W (υπερ-δειγματοληψία UHD). Ωστόσο, χρειάζεται τουλάχιστον τρεις τάξεις μεγέθους περισσότερο χρόνο, εμφανίζοντας συνδυασμένο δείκτη απόδοσης 0.091 Msamples/W. Μεταξύ των άλλων δύο επεξεργαστών, ο οκταπύρηνος *FX-8120* απαιτεί σχεδόν διπλάσια ισχύ σε σχέση με τον *i5-3470*: 97.5 W κατά την υπερ-δειγματοληψία εικόνων 1080p σε μέγεθος UHD, σε σύγκριση με τον τελευταίο ο οποίος καταναλώνει 50.78 W. Αντίστοιχα, ο συνδυασμένος δείκτης επιδόσεων τους αποτιμάται σε 0.042 και 0.112 Msamples/W.

Οι επεξεργαστές γραφικών εμφανίζουν παρόμοια συμπεριφορά κατανάλωσης, με τιμές που κυμαίνονται από 67 W (*GTX 960*) έως 76 W (*GTX 670*), με τη διαφορά ότι τελειώνουν τους υπολογισμούς σε δύο τάξεις μεγέθους λιγότερο χρόνο. Με συνδυασμένο δείκτη 21.66 Msamples/W, ο *GTX 960* συνιστά την πιο αποδοτική ενεργειακά λύση ακολουθούμενος από τον *GTX 670* με 20.69 και τον *GTX 550Ti* με 8.32 Msamples/W.

Όπως αναμενόταν, τα ολοκληρωμένα FPGA αποτελούν την πιο αποδοτική ενεργειακά λύση μεταξύ όλων των εξετασθέντων λύσεων επεξεργασίας. Παρατηρούμε ότι παρόλο που χαρακτηρίζονται από χρόνο υπολογισμού συγκρίσιμο με αυτόν των επεξεργαστών γραφικών, καταναλώνουν τουλάχιστον μία τάξη μεγέθους λιγότερη ισχύ. Καλούμε τον αναγνώστη επίσης να παρατηρήσει το αυξημένο διάστημα μεταξύ των επιδόσεων σε GPU και FPGA για εικόνες XE μικρού μεγέθους (αριστερό τμήμα της Εικ. 6.10), γεγονός το οποίο μπορεί να αποδοθεί στη χαμηλή αξιοποίηση των υπολογιστικών πόρων των GPU με συνέπεια την αδράνεια πολλών επεξεργαστικών πυρήνων. Επίσης, τα γραφήματα τείνουν να συγχλίνουν (βλ. δεξί τμήμα της Εικ. 6.10 και Πίνακα 6.4) καθώς οι σημαντικά αυξημένες απαιτήσεις μνήμης των εικόνων μεγαλύτερου μεγέθους οδηγούν σε οριακά επίπεδα τους διαθέσιμους πόρους μονάδων μνήμης των συστοιχιών FPGA, ενώ από την άλλη οι διατάξεις GPU μπορούν να διαθέτουν άπλετο χώρο μνήμης (π.χ. 4GB στη διάταξη υποδοχής του *GTX 960*). Όσον αφορά την από κοινού αποτίμηση επιδόσεων-ισχύος, τα ολοκληρωμένα FPGA χαμηλής αποδοτικότητας ξεχωρίζουν: η υλοποίηση 3 μηχανών στο xc7a100t-1 επιτυγχάνει τη μεγαλύτερη τιμή συνδυασμένου δείκτη ήτοι 2060.4 Msamples/W, ακολουθούμενη από το 7-485-2 (16) με 1460.37 Msamples/W και το xc6slx03-2 (12) με 1453.67 Msamples/W.

Εξετάζοντας την απόδοση από ενεργειακή σκοπιά, τα συγκεντρωτικά αποτελέσματα του πίνακα 6.5 μας βοηθούν να διακρίνουμε τη σχετική θέση κάθε πλατφόρμας κατά τρόπο αδιαμφισβήτητο. Έχοντας πλέον ως δείκτη τα Joules που απαιτούνται για τον υπολογισμό κάθε δείγματος ΥΕ, παρατηρούμε διαφορά δύο τάξεων μεγέθους μεταξύ CPU και GPU και μεταξύ GPU και FPGA. Η διακύμανση ενέργειας ελαττώνεται με την αύξηση του μεγέθους της εικόνας XE με συνέπεια να οδηγεί σε ντετερμινιστική συμπεριφορά και μικρότερες ενεργειακές απώλειες. Τέλος, είναι εμφανές ότι η ενεργειακή απόδοση τείνει να αυξάνεται σε ολοκληρωμένα κυκλώματα νεότερης γενιάς (όπως οι επεξεργαστές *i5-3470*, *GTX 960* και η σειρά 7 των συστοιχιών FPGA της Xilinx).

Πίνακας 6.5: Ενεργειακή αποτίμηση των υλοποιήσεων του αλγορίθμου *SIL-SEABI* σε CPU, GPU και FPGA.

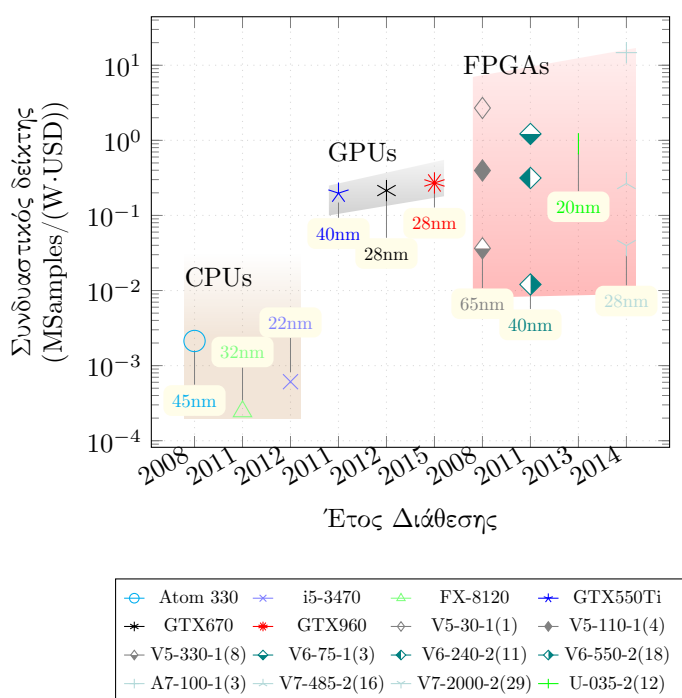
Πλατφόρμα	Ενέργεια Υλοποιήσεων SIL-SEABI (nJoules/δείγμα)						
	Μέγεθος Εικόνας ΥΕ ($f = 2$)						
	QCIF	CIF	SD1	720p	1080p	2160p	μ.τ.
Atom 330	38065	16679	11503	8021	8410	7660	15056
i5-3470	59899	20104	10013	7393	5696	5069	18029
FX-8120	97206	35224	22074	19490	19680	14456	34689
GTX550Ti	784.1	206.6	167.6	80.11	78.06	80.88	232.9
GTX670	865.1	245.2	104.2	53.12	38.18	37.94	223.9
GTX960	652.5	167.9	51.71	32.88	26.17	28.92	160.03
V5-30-1(1)	0.927	0.899	0.890	0.888	0.885	0.882	0.895
V5-110-1(4)	1.624	1.279	1.113	1.079	1.033	0.988	1.186
V5-330-1(8)	3.097	2.174	1.724	1.633	1.509	1.389	1.921
V6-75-1(3)	1.601	1.315	1.179	1.151	1.113	1.076	1.239
V6-240-2(11)	2.530	1.683	1.268	1.184	1.070	0.959	1.449
V6-550-2(18)	3.119	1.920	1.330	1.211	1.050	0.893	1.587
A7-100-1(3)	0.639	0.525	0.471	0.459	0.444	0.429	0.495
V7-485-2(16)	1.530	0.958	0.677	0.621	0.544	0.469	0.799
V7-2K-2(29)	2.957	1.709	1.095	0.970	0.803	0.640	1.362
U-035-2(12)	1.379	0.904	0.671	0.624	0.560	0.498	0.773

6.3.6 Σύγκρισεις πλατφόρμας υλοποίησης: κόστος ανάπτυξης

Στον κατακόρυφο άξονα της εικόνας 6.11 παρατίθεται επίσης το μέσο κόστος κάθε ολοκληρωμένου κυκλώματος σε δολάρια (USD) για ποσότητες 1000 τεμαχίων όπως αυτό κατεγράφη κατά τη διάρκεια συγγραφής της εργασίας¹. Είναι εμφανές πως τα ολοκληρωμένα FPGA διέπονται γενικά από το υψηλότερο κόστος κτήσης: παρόλο που για το χαμηλής κατανάλωσης A7-100-1 το κόστος δεν υπερβαίνει τα \$140, σε μέσης αποδοτικότητας ολοκληρωμένα όπως τα V5-110-1 και V6-240-2 μπορεί να ανέλθει σε περίπου \$2300, φτάνοντας τα \$6206 και \$23838 για τα υψηλής αποδοτικότητας V6-550-2 και V7-2000-2 αντίστοιχως. Λαμβάνοντας υπόψιν στην ανάλυση και το κόστος κτήσης ο αναγνώστης μπορεί να αποκτήσει μια πιο ολοκληρωμένη εικόνα. Για παράδειγμα, η διάταξη των 16 παράλληλων μηχανών στο 7-485-2 συγκρινόμενη με τη διάταξη γραφικών GTX 670 εκτός από αυξημένη ενεργειακή αποδοτικότητα εμφανίζει και ανάλογα αυξημένο κόστος κτήσης, οδηγώντας σε παρόμοια τιμή συνδυαστικού δείκτη απόδοσης-ισχύος-κόστους με τιμή 0.2 Msamples/(W·USD). Οι υψηλότερες τιμές του συνδυαστικού αυτού δείκτη παρατηρούνται σε χαμηλής ισχύος διατάξεις FPGA όπως η περίπτωση των τριών μηχανών στο 7-100-1 και της μίας μηχανής στο 5-30-1 με αντίστοιχες τιμές 14.7 και 2.7 Msamples/(W·USD).

¹Όπως συλλέχθηκαν διαδικτυακά κυρίως μέσω της μηχανής αναζήτησης <https://octopart.com> (Απρίλιος 2016).

SIL-SEABI - Χρόνος, Ισχύς και Κόστος Κτήσης (1080p → 2160p)



Εικόνα 6.12: Συνδυαστικός δείκτης απόδοσης-κόστους του αλγορίθμου SIL-SEABI σε διάφορες πλατφόρμες υλοποίησης.

Η συγκέντρωση όλων των αποτελεσμάτων στο γράφημα της Εικ. 6.12 οδηγεί στην παρατήρηση ότι ο συνδυαστικός δείκτης απόδοσης τείνει να αυξάνεται σε όλες τις πλατφόρμες όσο παράλληλα μειώνεται και το μέγεθος των ημιαγωγικών διατάξεων (transistors) σε βάθος χρόνου. Όσον αφορά τις πλατφόρμες συγκριτικά, ο συνδυαστικός δείκτης των διατάξεων γραφικών είναι άνω των δύο τάξεων μεγέθους μεγαλύτερος αυτού των επεξεργαστών γενικού σκοπού. Επίσης, οι διατάξεις FPGA διατηρούν το πλεονέκτημα απόδοσης έναντι όλων των άλλων υλοποιήσεων ακόμη και όταν στην αποτίμηση συμπεριληφθεί επιπλέον το κόστος κτήσης. Από την άλλη πλευρά βέβαια, οι λύσεις οι οποίες βασίζονται σε FPGA υψηλής αποδοτικότητας απαιτούν ένα τίμημα εξαιρετικά δυσανάλογο - δεδομένων των ουσιαστικών πρόσθετων επιδόσεων τις οποίες προσφέρουν σε σχέση με τις λύσεις μέσης αποδοτικότητας. Αξιοσημείωτο επίσης είναι το γεγονός ότι οι διατάξεις γραφικών τείνουν να προσφέρουν το ίδιο επίπεδο απόδοσης με τις χαμηλής-προς μέσης αποδοτικότητας διατάξεις FPGA καθώς επίσης και να προσεγγίζουν το συνδυαστικό δείκτη απόδοσης των διατάξεων FPGA μέσης αποδοτικότητας. Ταυτόχρονα, σε ολοκληρωμένα κυκλώματα FPGA χαμηλής ισχύος παρατηρείται τόσο αύξηση στους διαθέσιμους πόρους, όσο και ελάττωση του κόστους κτήσης (Εικ. 6.12).

Διακριτός είναι επίσης και ο συνολικός χρόνος σχεδίασης και ανάπτυξης ανά πλατφόρμα. Υποθέτοντας ότι η ανάπτυξη ξεκινά από την αρχή, οι υλοποιήσεις σε επεξεργαστές γραφικών απαιτούν έναν ανθρωπο-μήνα ανά αλγόριθμο συμπεριλαμβανομένων των βελτιστοποιήσεων, τη στιγμή που οι αντίστοιχες υλοποιήσεις σε FPGA χρειάζονται περίπου τριπλάσια προσπάθεια. Συμπερασματικά, η αποτίμηση των συγκεκριμένων αλγορίθμων και των αντίστοιχων υλοποιήσεών τους οδηγεί στο συμπέρασμα ότι οι διατάξεις FPGA είναι κατά μία τάξη μεγέθους ταχύτερες από τις διατάξεις γραφικών. Εξετάζοντας δε από κοινού επιδόσεις και κατανάλωση, παρατηρήσαμε ότι οι χαμηλής ολοκλήρωσης FPGA διατάξεις μπορούν να είναι ενεργειακά έως και δύο τάξεις μεγέθους αποδοτικότερες από τους υψηλής αποδοτικότητας επεξεργαστές γραφικών. Η ίδια αναλογία επιδόσεων παρατηρείται και μεταξύ των τελευταίων και των πολυπύρηνων επεξεργαστών γενικού σκοπού. Τέλος, όταν στην αποτίμηση ληφθεί υπόψιν και το κόστος κτήσης για κάθε ολοκληρωμένο κύκλωμα, τα αποτελέσματα οδηγούν στη διαπίστωση ότι οι επεξεργαστές γραφικών και οι χαμηλής ολοκλήρωσης FPGA διατάξεις συνιστούν τις προτεινόμενες πλατφόρμες ανάπτυξης.

Πλεονεκτήματα και εφαρμογές των προτεινόμενων τεχνικών

Στα προηγούμενα κεφάλαια περιγράφηκαν: α) η συμβολή του αλγορίθμου ως μία αυτόνομη λύση για τη βελτίωση της ποιότητας υπερανάλυσης χαμηλής πολυπλοκότητας και β) οι τεχνικές υλοποίησης σε μηχανές παράλληλης επεξεργασίας μέσω των οποίων αξιοποιείται η χαμηλή αυτή πολυπλοκότητα για την παροχή λύσεων υπερ-ανάλυσης σε πραγματικό χρόνο. Σε αυτό το κεφάλαιο θα παρουσιαστούν τα αποτελέσματα εφαρμογής του αλγορίθμου ως επικουρικής λύσης για τη βελτίωση της ποιότητας ΥΑ περίπλοκων υπερσύγχρονων μεθόδων τεχνολογικής αιχμής. Τέλος, θα εξεταστεί η εφαρμογή των τεχνικών επιτάχυνσης σε μεθόδους στερεοσκοπικής αντιστοιχίας, προκειμένου να αποτιμηθεί η αποτελεσματικότητά τους σε ένα απαιτητικότερο σενάριο επεξεργασίας εικόνας.

7.1 Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων αλγορίθμων ΥΑ

Στο κεφάλαιο 5 εισαγάγαμε τον αλγόριθμο ΥΑ *L-SEABI*, ο οποίος επιτυγχάνει συγκρίσιμα αποτελέσματα ποιότητας με εκείνα των εξαιρετικά περίπλοκων μεθόδων ΥΑ έχοντας σημαντικά χαμηλότερο υπολογιστικό κόστος (3 τάξεις μεγέθους χαμηλότερο χρόνο εκτέλεσης). Ο αλγόριθμος αποτελείται από τη φάση δόμησης η οποία εκτελείται μία φορά και τη φάση βελτίωσης η οποία είναι επαναληπτική. Εξετάσαμε επίσης τη μέθοδο *SIL-SEABI*, την υποπερίπτωση μίας επανάληψης του αλγορίθμου *L-SEABI* η οποία μειώνει περαιτέρω το χρόνο εκτέλεσης κατά μία τάξη μεγέθους με συγκριτικά αμελητέα υποβάθμιση ποιότητας. Η σύγκριση διεξήχθη με μεθόδους ΥΑ τεχνολογικής αιχμής: α) λεξικοπαγείς, όπως η αραιή αναπαράσταση με υπερ-πλήρη λεξικά [98], η προσαρμοστική επιλογή/κανονικοποίηση του χώρου αραιής αναπαράστασης [23] (*ASDS-AR-NL*) και η Αγκυροβολημένη Παλινδρόμηση Γειτονίας (*ANR*) [86], β) μεθόδους *EOIP*, όπως η μη τοπική *EOIP* [22] (*NLIBP*) και γ) η ΥΑ πολλαπλών εικόνων εισόδου με Bayesian *πρότερη γνώση* [90]. Πρόσφατα, οι συγγραφείς του [24] πρότειναν την έννοια της μη τοπικής αυτοπαλινδρομικής μοντελοποίησης (*Non-local Autoregressive Modeling Narm*) δηλαδή την προσαρμοστική εκμετάλλευση των

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων
 ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

Πίνακας 7.1: Αντικειμενική σύγκριση ποιότητας μεταξύ του αλγορίθμου *L-SEABI* της εκτέλεσης μίας επανάληψης αυτού, της φάσης δόμησης αυτού και των μεθόδων ΥΑ *NARM* και της ΥΑ μέσω παραμορφώσιμων πλακιδίων (παράγοντας υπερ-δειγματοληψίας $f=2$). Όλοι οι αλγόριθμοι εκτελούνται με τις προτεινόμενες παραμέτρους από τους αντίστοιχους συγγραφείς. Χαμηλότερες τιμές της μετρικής BRISQUE υποδηλώνουν υψηλότερη αντικειμενική ποιότητα εικόνων.

Μέθοδος:	L-	SIL-	L-	NARM	DPSR	
ανάλυση	SEABI	SEABI	SEAI	[24]	[101]	
MSSIM	QCIF	0.9114	0.9086	0.9013	0.9278	0.9106
	CIF	0.8763	0.8743	0.8604	0.8817	0.8834
	SD1	0.9702	0.9679	0.9652	0.9730	0.9751
	720p	0.9403	0.9567	0.9537	0.9618	0.9636
	1080p	0.9796	0.9784	0.9783	0.9808	0.9816
	2160p	0.9880	0.9879	0.9874	0.9887	0.9883
μ.τ.	0.9474	0.9456	0.9411	0.9523	0.9504	
BRISQUE	QCIF	39.9237	43.5880	42.5035	39.7551	30.1888
	CIF	33.8815	36.1080	33.0046	45.2763	42.7670
	SD1	27.8568	29.5815	30.4322	35.6517	31.4534
	720p	41.6382	43.5580	46.1487	46.9208	41.8770
	1080p	40.4554	42.0595	46.7294	47.1687	44.9506
	2160p	56.8680	57.9444	58.4345	59.8403	55.7759
μ.τ.	40.1054	42.1399	42.8755	45.7688	41.1964	

μη τοπικών πλεονασμών μίας εικόνας εντός ενός πλαισίου αραιής αναπαράστασης βελτιώνοντας έτσι την απόδοση του τυποποιημένου μοντέλου αραιής αναπαράστασης. Οι συγγραφείς του [101] βελτίωσαν την προσέγγιση αραιής αναπαράστασης προτείνοντας τη χρήση πλακιδίων/πεδίων κανονικοποιημένης παραμόρφωσης αντί για σταθερά ανυσματικά πλακίδια (ΥΑ Παραμορφώσιμων Πλακιδίων, *DPSR*). Κοινός τόπος σε όλες τις προαναφερθείσες μεθόδους είναι το αρχικό βήμα ανακατασκευής μέσω του οποίου δομείται μια αρχική εκτίμηση της εικόνας μέσω δικυβικής παρεμβολής.

Δεδομένου ότι ο αλγόριθμος *L-SEABI* παρέχει βελτιωμένη ποιότητα σε σχέση με τις κοινές μεθόδους παρεμβολής θα διερευνήσουμε το ενδεχόμενο χρήσης του *L-SEABI* για τη βελτίωση της απόδοσης των περίπλοκων μεθόδων ΥΑ. Συνεπώς, στην ενότητα αυτή θα εφαρμοστεί α) ο *L-SEABI*, β) ο *SIL-SEABI* και γ) η φάση δόμησης των α), β) (στην οποία θα αναφερόμαστε από εδώ και στο εξής ως *L-SEAI*), στη θέση του αρχικού σταδίου αναδόμησης των μεθόδων [98, 23, 86, 22, 101, 24] κατά την υπερ-δειγματοληψία κάθε εικόνας ΧΕ κατά συντελεστή δύο σε κάθε διάσταση. Για τη μέτρηση της ποιότητας του αποτελέσματος, θα χρησιμοποιηθεί η μέθοδος MSSIM [93] η οποία χρησιμοποιεί την αρχική εικόνα ως αναφορά και η μέθοδος BRISQUE [57] η οποία εφαρμόζεται χωρίς εικόνα αναφοράς.

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων
 ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

Πίνακας 7.2: Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f=2$) όταν χρησιμοποιούνται οι αλγόριθμοι *L-SEABI* (a), *SIL-SEABI* (b) και *L-SEAI* (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική MSSIM).

	↓Μετρική	Αλγόριθμος→	NLIBP	ANR	Yang et al.
	ανάλυση↓	Αρχικοποίηση	[22]	[86]	[98]
ΔMSSIM _(επιλειτουργισμένο-τυπικό)	QCIF	(a)	-0.13793	-0.02854	-0.02495
		(b)	-0.14526	-0.03719	-0.03268
		(c)	-0.04064	-0.00560	-0.00984
	CIF	(a)	-0.13655	-0.02631	-0.01849
		(b)	-0.13367	-0.03043	-0.02266
		(c)	-0.03763	-0.00423	-0.01087
	SD1	(a)	-0.06806	-0.01726	-0.01665
		(b)	-0.06947	-0.01901	-0.01864
		(c)	-0.02102	-0.00902	-0.01039
	720p	(a)	-0.07074	-0.01943	-0.01730
		(b)	-0.07234	-0.02162	-0.01914
		(c)	-0.01511	-0.01082	-0.01407
	1080p	(a)	-0.05187	-0.01130	-0.01427
		(b)	-0.05229	-0.01223	-0.01473
		(c)	-0.02102	-0.00564	-0.00973
	2160p	(a)	-0.02498	-0.05470	-0.01093
		(b)	-0.02561	-0.00592	-0.01102
		(c)	-0.00753	-0.00282	-0.00866
	avg	(a)	-0.08169	-0.01805	-0.01710
		(b)	-0.08311	-0.02106	-0.01981
		(c)	-0.02379	-0.00636	-0.00106

Ο Πίνακας 7.1 παρουσιάζει τη μέση τιμή των μετρικών MSSIM και BRISQUE για τις μεθόδους *L-SEABI*, *SIL-SEABI*, *L-SEAI*, εν συγκρίσει με τις *Narm* και *DPSR* για το σύνολο των δεδομένων του κεφαλαίου 5 (26 εικόνες με εύρος αναλύσεων από 176×144 έως 3840×2160 εικονοστοιχεία). Τα αποτελέσματα καταδεικνύουν ότι η προσάρτηση επαναληπτικών βελτιώσεων στη μέθοδο *L-SEAI* αυξάνει οριακά την ποιότητα του αποτελέσματος όπως αξιολογείται με την MSSIM μετρική, αν και η παρατηρηθείσα διαφορά καταγράφεται πιο έντονα μέσω της μετρικής BRISQUE. Επιπλέον, όλες οι υποπεριπτώσεις του αλγορίθμου *L-SEABI* αποδίδουν τουλάχιστον συγκρίσιμη ποιότητα (μετρική MSSIM) ή ακόμη και ανώτερη ποιότητα (μετρική BRISQUE) σε σχέση με εκείνη των μεθόδων *Narm* και *DPSR* (μετρική MSSIM). Τα αποτελέσματα αυτά συμφωνούν με εκείνα που παρουσιάστηκαν στο Κεφ. 5 όσον αφορά τις λεξικοπαγείς μεθόδους ΥΑ με μία εικόνα αναφοράς.

Οι Πίνακες 7.2, 7.3, 7.4 και 7.5 παρουσιάζουν τη μέση τιμή της διαφοράς

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων
 ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

Πίνακας 7.3: Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f=2$) όταν χρησιμοποιούνται οι αλγόριθμοι *L-SEABI* (a), *SIL-SEABI* (b) και *L-SEAI* (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική MSSIM).

	↓Μετρική	Αλγόριθμος→	NARM	DPSR	ASDS-AR-NL
	ανάλυση↓	Αρχικοποίηση	[24]	[101]	[23]
Δ MSSIM _(enhanced-original)	QCIF	(a)	0.00337	0.01224	0.00242
		(b)	0.00321	0.01208	0.00223
		(c)	0.00469	0.00918	0.00178
	CIF	(a)	0.00220	0.01481	-0.00021
		(b)	0.00208	0.01444	-0.00011
		(c)	0.00235	0.01121	0.00138
	SD1	(a)	0.00116	0.00225	0.00160
		(b)	0.00119	0.00214	0.00155
		(c)	0.00113	0.00164	0.00150
	720p	(a)	0.00116	0.00276	0.00171
		(b)	0.00113	0.00267	0.00169
		(c)	0.00114	0.00206	0.00164
	1080p	(a)	0.00128	0.00087	0.00168
		(b)	0.00126	0.00083	0.00168
		(c)	0.00125	0.00064	0.00165
	2160p	(a)	-0.00005	0.00037	-
		(b)	-0.00004	0.00035	-
		(c)	-0.00004	0.00029	-
	avg	(a)	0.00152	0.00555	-
		(b)	0.00147	0.00542	-
		(c)	0.00175	0.00417	-

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων
 ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

Πίνακας 7.4: Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f=2$) όταν χρησιμοποιούνται οι αλγόριθμοι *L-SEABI* (a), *SIL-SEABI* (b) και *L-SEAI* (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική BRISQUE). Χαμηλότερες τιμές διαφοράς Δ BRISQUE υποδηλώνουν υψηλότερη αντικειμενική ποιότητα εικόνων.

	↓Μετρική ανάλυση↓	Αλγόριθμος→ Αρχικοποίηση	NLIBP	ANR	Yang et al.
			[22]	[86]	[98]
Δ BRISQUE _(enhanced-original)	QCIF	(a)	-7.41320	-3.68780	-3.47380
		(b)	-2.62820	-0.57520	-3.28160
		(c)	-6.34600	0.55840	2.81730
	CIF	(a)	7.80928	-2.73474	-5.99424
		(b)	9.06760	-2.09434	-5.41352
		(c)	0.66116	-10.2587	-4.09130
	SD1	(a)	-1.29152	-3.68073	-6.15684
		(b)	3.19676	-2.51239	-5.29779
		(c)	-2.40543	-4.19604	-0.94454
	720p	(a)	-8.77396	-2.94443	-6.10462
		(b)	-5.48932	-2.42523	-4.26859
		(c)	-5.10736	-0.39096	-0.25632
	1080p	(a)	-7.86562	-7.99265	-10.8402
		(b)	-4.88077	-7.44972	-9.51325
		(c)	-3.97212	-2.69070	-3.31287
	2160p	(a)	-14.1632	0.58664	-5.16501
		(b)	-11.7626	1.43108	-3.87191
		(c)	-7.66023	1.50892	-1.82637
	avg	(a)	-5.2830	-3.40895	-6.28912
		(b)	-2.0827	-2.27096	-5.27444
		(c)	-4.13833	-2.57818	-1.26902

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων
 ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

Πίνακας 7.5: Αντικειμενική σύγκριση υπερσύγχρονων αλγορίθμων ΥΑ (παράγοντας υπερ-δειγματοληψίας $f=2$) όταν χρησιμοποιούνται οι αλγόριθμοι *L-SEABI* (a), *SIL-SEABI* (b) και *L-SEAI* (c) στο αρχικό στάδιο ανακατασκευής της εικόνας ΥΕ έναντι των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς (μετρική BRISQUE). Χαμηλότερες τιμές διαφοράς Δ BRISQUE υποδηλώνουν υψηλότερη αντικειμενική ποιότητα εικόνων.




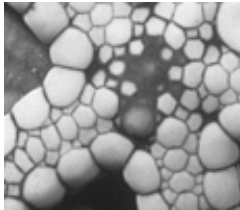



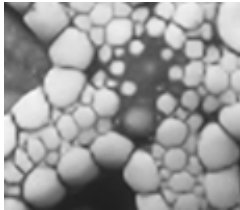



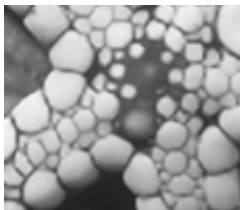

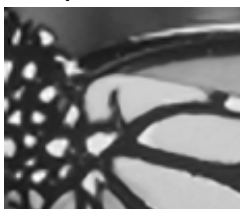

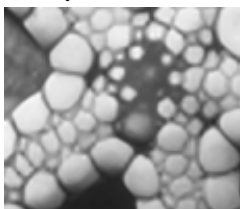


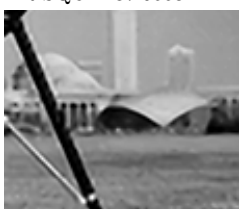
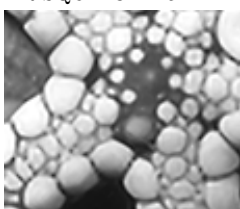
	↓Μετρική	Αλγόριθμος→	NARM	DPSR	ASDS-AR-NL
	ανάλυση↓	Αρχικοποίηση	[24]	[101]	[23]
Δ BRISQUE _(enhanced-original)	QCIF	(a)	-0.64420	-7.93090	-0.61830
		(b)	-2.81310	-7.48240	-0.55620
		(c)	-2.56050	-7.89660	-0.12090
	CIF	(a)	-0.42284	-3.12196	-0.17576
		(b)	-0.15354	-2.88236	-0.03572
		(c)	-0.36940	-3.10202	0.08386
	SD1	(a)	-0.72980	-2.02076	-0.30873
		(b)	-0.44140	-2.00466	-0.22110
		(c)	-0.61230	-2.17134	-0.06533
	720p	(a)	0.05638	-0.93106	-0.0634
		(b)	-0.38838	-0.83522	-0.07618
		(c)	0.09672	-0.79050	-0.03506
	1080p	(a)	1.36240	-1.83855	-0.01110
		(b)	0.48452	-1.76727	0.00032
		(c)	0.70232	-1.56985	-0.00012
	2160p	(a)	-0.40221	-0.40991	-
		(b)	-0.48665	-0.33370	-
		(c)	-0.17654	-0.28128	-
	avg	(a)	-0.13004	-2.70886	-
		(b)	-0.63309	-2.55094	-
		(c)	-0.48661	-2.63526	-

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

ποιότητας μεταξύ των αποτελεσμάτων που λαμβάνονται από τις μεθόδους [98, 23, 86, 22, 101, 24] όταν η αρχική ανακατασκευή διενεργείται με μία εκ των *L-SEABI* / *SIL-SEABI* / *L-SEAI* και τα αποτελέσματα που λαμβάνονται με τις ίδιες μεθόδους αλλά με τη χρήση των παραμέτρων που προτείνονται από τους αντίστοιχους συγγραφείς τους κατά τη φάση δόμησης. Βάσει των αποτελεσμάτων, αποδεικνύεται ότι η χρήση εναλλακτικής ανακατασκευής για τη μέθοδο [22] οδηγεί σε ελαφρά υποβάθμιση της μετρικής *MSSIM* κατά 0,08 κατά μέσο όρο, η οποία καθίσταται εμφανέστερη σε εικόνες μεγέθους 352×288 δειγμάτων. Από την άλλη πλευρά, η διαφορά *ΔBRISQUE* καταγράφει αύξηση της ποιότητας, ειδικά όταν γίνεται χρήση των *L-SEABI* / *SIL-SEABI* σε 720p ή υψηλότερες αναλύσεις. Να σημειωθεί εδώ ότι η μέθοδος [22] όπως προτείνεται από τους συγγραφείς εκτελεί 20 επαναλήψεις οπισθοπροβολής, και, όταν συνδυάζεται με τα αποτελέσματα των *L-SEABI* / *SIL-SEABI* τείνει να παραμορφώνει το αποτέλεσμα. Ο λόγος είναι ότι οι μέθοδοι αυτές εμπεριέχουν ήδη το στάδιο της οπισθοπροβολής με αποτέλεσμα να υπερενισχύεται τελικώς το υψίσυχο περιεχόμενο της εικόνας. Συνεπώς, η μέθοδος *NLIBP*, παράγει πιο ισορροπημένο αποτέλεσμα χρησιμοποιώντας την υποπερίπτωση *L-SEAI* σε όλες τις αναλύσεις, ενώ οι *L-SEABI* και *SIL-SEABI* θα μπορούσαν να αξιοποιηθούν περισσότερο σε περιεχόμενο υψηλής ευκρίνειας. Παρόμοια αποτελέσματα ελήφθησαν και κατά την εξέταση της μεθόδου *ANR*: καταγράφεται μια μέση υποβάθμιση της μετρικής *MSSIM* κατά 0,02 ενώ η *BRISQUE* ευνοεί την εφαρμογή των *L-SEABI* / *SIL-SEABI*, εμφανίζοντας αύξηση της ποιότητας σε όλες τις αναλύσεις με εξαίρεση την ανάλυση 354×288 - όπου η *L-SEAI* παρέχει καλύτερα αποτελέσματα - και την ανάλυση 3840×2160 , όπου η ποιότητα φθίνει οριακά. Όταν ενισχύουμε τη μέθοδο [98] παρατηρούμε παρόμοια συμπεριφορά αποτελεσμάτων *MSSIM* με εκείνη της μεθόδου *ANR*. Σε αυτή την περίπτωση, οι μέθοδοι *L-SEABI* και *SIL-SEABI* αυξάνουν την ποιότητα του αποτελέσματος σε όλες τις αναλύσεις βάσει της μετρικής *BRISQUE*. Όσον αφορά τις μεθόδους *Narm* και *ASDS-AR-NL*, η μετρική *MSSIM* εμφανίζει συνολικά μια αμελητέα αύξηση κατά περίπου 0.002, ενώ η μετρική *BRISQUE* καταδεικνύει ότι οι *SIL-SEABI* και *L-SEABI* παρέχουν αντιστοίχως την υψηλότερη ποιότητα σε όλες τις αναλύσεις. Οι προτεινόμενες μέθοδοι χαμηλής πολυπλοκότητας ενισχύουν επίσης την ποιότητα του αλγορίθμου *DPSR*, κατά 0,005 και 2,63 όσον αφορά τις μετρικές *MSSIM* και *BRISQUE* αντιστοίχως. Ως εκ τούτου, τα παραπάνω αποτελέσματα δείχνουν ότι στις περισσότερες περιπτώσεις ο αλγόριθμος *SIL-SEABI* μπορεί όντως να χρησιμοποιηθεί για την περαιτέρω βελτίωση των αποτελεσμάτων ΥΑ των πιο εξελιγμένων τεχνολογικά μεθόδων, όπως αυτή αξιολογείται αντικειμενικά από τις μετρικές *MSSIM* και *BRISQUE*.

Για την υποκειμενική αξιολόγηση των αποτελεσμάτων κατόπιν της επικουρικής εφαρμογής των μεθόδων μας, στο Σχ. 7.1 παρουσιάζονται οι εικόνες ΥΕ όπως προέκυψαν από την εφαρμογή των [24] και [98] κατά την επεξεργασία των εικόνων *Carphone* ανάλυσης 176×144 και *Butterfly*, *Cameraman*, *Starfish* ανάλυσης 256×256 . Από τις παραχθείσες εικόνες φαίνεται ότι η χρήση του *SIL-SEABI* για τη βελτίωση της μεθόδου *Narm*, (Εικόνα 7.1θ', 7.1ι', 7.1ια' και 7.1ιβ') ελαττώνει αισθητά το φαινόμενο οδόντωσης (εικόνες *Carphone*, *Cameraman* και *Starfish*).

7.1. Ο αλγόριθμος *L-SEABI* ως μηχανισμός βελτίωσης των υπερσύγχρονων
 ΚΕΦΑΛΑΙΟ 7 αλγορίθμων ΥΑ

			
(α') Carphone: αρχική εικόνα	(β') Butterfly: αρχική εικόνα	(γ') Cameraman: αρχική εικόνα	(δ') Starfish: αρχική εικόνα
			
(ε') τυπική ΥΑ: [24] MSSIM: 0.92782 BRISQUE: 39.7551	(στ') τυπική ΥΑ: [24] MSSIM: 0.925027 BRISQUE: 33.8404	(ζ') τυπική ΥΑ: [24] MSSIM: 0.865780 BRISQUE: 39.3129	(η') τυπική ΥΑ: [24] MSSIM: 0.914907 BRISQUE: 30.1323
			
(θ') [24]+ <i>SIL-SEABI</i> MSSIM: 0.93103 BRISQUE: 39.1109	(ι') [24]+ <i>SIL-SEABI</i> MSSIM: 0.926914 BRISQUE: 32.3599	(ια') [24]+ <i>SIL-SEABI</i> MSSIM: 0.869826 BRISQUE: 38.4339	(ιβ') [24]+ <i>SIL-SEABI</i> MSSIM: 0.915686 BRISQUE: 29.8780
			
(ιγ') τυπική ΥΑ: [98] MSSIM: 0.91694 BRISQUE: 31.6656	(ιδ') τυπική ΥΑ: [98] MSSIM: 0.909051 BRISQUE: 31.0606	(ιε') τυπική ΥΑ: [98] MSSIM: 0.863308 BRISQUE: 37.0668	(ιστ') τυπική ΥΑ: [98] MSSIM: 0.904035 BRISQUE: 34.1451
			
(ιζ') [98]+ <i>SIL-SEABI</i> MSSIM: 0.88426 BRISQUE: 28.3840	(ιη') [98]+ <i>SIL-SEABI</i> MSSIM: 0.860352 BRISQUE: 26.8966	(ιθ') [98]+ <i>SIL-SEABI</i> MSSIM: 0.837666 BRISQUE: 33.3080	(ιχ') [98]+ <i>SIL-SEABI</i> MSSIM: 0.879176 BRISQUE: 29.1027

Εικόνα 7.1: Υποκειμενική σύγκριση ποιότητας των μεθόδων [24], [98]: τυπική επεξεργασία έναντι εμπλουτισμένης με τον αλγόριθμο *SIL-SEABI* ($f = 2$).

7.2. Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγόριθμους ΚΕΦΑΛΑΙΟ 7 στερεοσκοπικής αντιστοίχισης

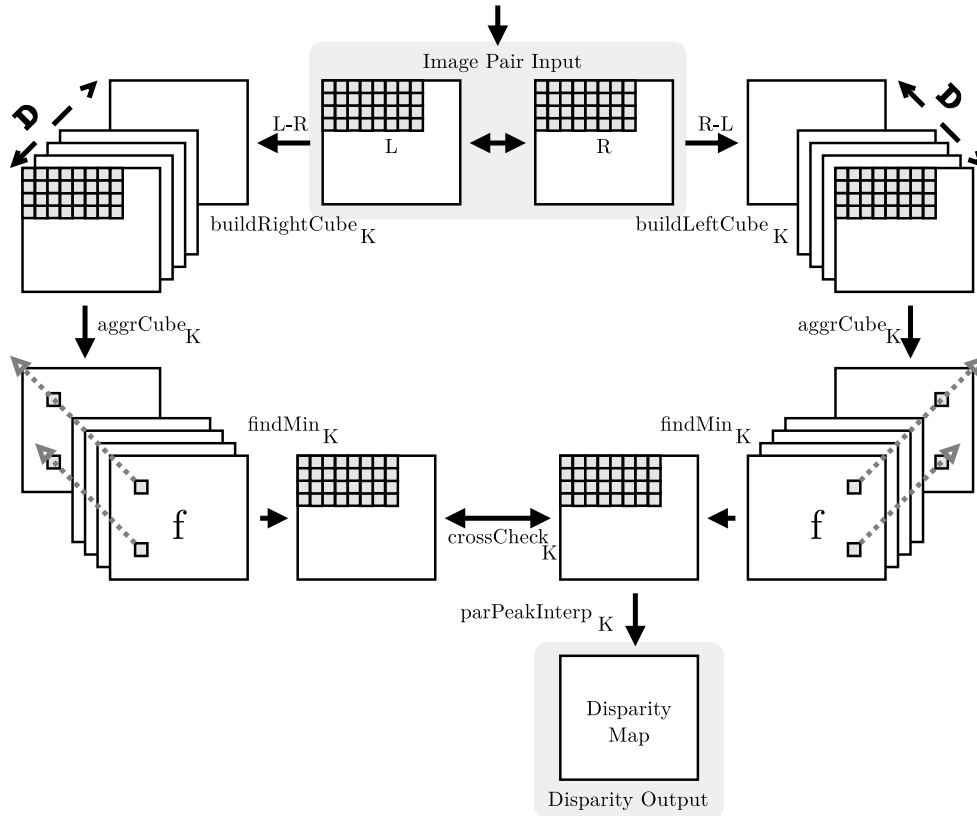
Επίσης, όταν χρησιμοποιείται ο *SIL-SEABI* κατά τη φάση ανακατασκευής του [98] (Εικόνες 7.1ιζ', 7.1ιη', 7.1ιθ' και 7.1ιχ') παρατηρείται αισθητή βελτίωση αντίθεσης στο οπτικό αποτέλεσμα. Παρόμοια συμπεριφορά (δηλαδή ελάττωση οδόντωσης και βελτίωση αντίθεσης) παρατηρείται και όταν ο *SIL-SEABI* εφαρμόζεται αντιστοίχως πριν από τις μεθόδους *DPSR* [101] και *ANR* [86].

7.2 Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγόριθμους στερεοσκοπικής αντιστοίχισης

Για τη βελτίωση της μελέτης αξιολόγησης και την επιβεβαίωση των συμπερασμάτων της ανάλυσης του αλγορίθμου *L-SEABI*, εφαρμόζουμε εν συνεχεία τις προτεινόμενες τεχνικές σε μια διαφορετική κατηγορία αλγορίθμων υπολογιστικής όρασης. Εξετάζουμε το πρόβλημα της στερεοσκοπικής αντιστοίχισης, η οποία σε γενικές γραμμές είναι υπολογιστικά απαιτητικότερη της ΥΑ. Συνήθως, οι αλγόριθμοι στερεοσκοπικής αντιστοίχισης βασίζουν την εκτέλεσή τους σε απαιτητικές διεργασίες σύγκρισης μεταξύ δύο εικόνων μέσω μίας επαναληπτικής προσέγγισης πλήρους αναζήτησης. Αυτό το χαρακτηριστικό μας επιτρέπει να αξιολογήσουμε τόσο τις τεχνικές παραλληλισμού όσο και τις πλατφόρμες υλοποίησης σε ένα αλγοριθμικό σενάριο το οποίο απαιτεί ενισχυμένη προσωρινή αποθήκευση και επαναχρησιμοποίηση των δεδομένων. Επιπλέον, είναι λιγότερο πρόσφορο για υπολογισμούς κατά τη διάρκεια μετάδοσης των εικόνων λόγω της περιορισμένης χωρικής τοπικότητας την οποία παρουσιάζει. Στην τρέχουσα ενότητα εξετάζουμε τον πιο αντιπροσωπευτικό και γνωστό αλγόριθμο της κατηγορίας, [84], τον αλγόριθμο ανομοιότητας.

Ο εξεταζόμενος αλγόριθμος ανομοιότητας επεξεργάζεται μία στερεοσκοπική εικόνα αποτελούμενη από $2 \times W \times H$ 8-bit δείγματα συνολικά, και εξαγάγει ένα χάρτη ανομοιότητας μεγέθους $W \times H$ με ακρίβεια ενός τετάρτου του εικονοστοιχείου [51]. Εσωτερικά, παράγονται δύο χάρτες ανομοιότητας διαστάσεων $W \times H$ ο ένας μετά τον άλλο (ο πρώτος βασίζεται στην αριστερή εικόνα και ο δεύτερος στη δεξιά εικόνα του στερεοσκοπικού ζεύγους). Τα δεδομένα από τους δύο χάρτες διασταυρώνονται ώστε να διατηρηθούν μόνο οι ανομοιότητες οι οποίες παρουσιάζονται και στους δύο χάρτες. Για την παραγωγή του αριστερού χάρτη, ο αλγόριθμος επεξεργάζεται τις διορθωμένες εικόνες επαναληπτικά με τη μέθοδο του δυναμικού προγραμματισμού. Σε κάθε επανάληψη, η δεξιά εικόνα ολισθαίνει κατά 1 εικονοστοιχείο στην οριζόντια διεύθυνση και αφαιρείται από την αριστερή εικόνα, παράγοντας $W \times H$ τιμές διαφορών οι οποίες αντιστοιχούν σε ένα επίπεδο του χώρου ανομοιότητας. Μετά από D επαναλήψεις, όπου D το πλήθος των εξεταζόμενων ανομοιοτήτων, ο αλγόριθμος ολοκληρώνει τη σύνθεση του χώρου ανομοιότητας αποθηκεύοντας $D \times W \times H$ τιμές. Εν συνεχεία οι τιμές αυτές αθροίζονται για κάθε εικονοστοιχείο και κάθε ένα από τα D επίπεδα με την ολίσθηση ενός παραθύρου 7×7 εικονοστοιχείων και τον υπολογισμό ενός αθροίσματος Γκαουσιανών βαρών για κάθε παράθυρο διαφορών 7×7 εικονοστοιχείων.

7.2. Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγόριθμους
 ΚΕΦΑΛΑΙΟ 7 σtereοσκοπικής αντιστοίχισης



Εικόνα 7.2: Αλγόριθμος Ανομοιότητας στην πλατφόρμα CUDA: ροή επεξεργασίας στη ΜΕΓ και πυρήνες επεξεργασίας.

Κάθε άθροισμα αντιστοιχεί σε μία μετρική ομοιότητας μεταξύ μιας περιοχής στην αριστερή και τη δεξιά εικόνα, δηλαδή παρέχει έναν τρόπο να συμπεράνουμε τη μετατόπιση ενός εικονοστοιχείου της αριστερής εικόνας στο επίπεδο της δεξιάς εικόνας. Πιο συγκεκριμένα, για κάθε εικονοστοιχείο της αριστερής εικόνας ο αλγόριθμος επιλέγει την ελάχιστη τιμή από τις D συνολικά διαφορές που υπολογίστηκαν κατά μήκος της ακτίνας βάθους (κατά μήκος της επιπολικής γραμμής του εικονοστοιχείου, ή πιο απλά ανάμεσα στα D επίπεδα του χώρου ανομοιότητας). Η τιμή αυτή αποτελεί και το αποτέλεσμα ανομοιότητας του εικονοστοιχείου. Τέλος, για την επίτευξη κλασματικής ακρίβειας ο αλγόριθμος επανυπολογίζει κάθε τιμή ανομοιότητας μέσω παραβολικής παρεμβολής μεταξύ των τριών τιμών γύρω από το ελάχιστο (πάντοτε κατά μήκος της ακτίνας βάθους). Ο δεξιός χάρτης βάθους παράγεται κατά παρόμοιο τρόπο με ολίσθηση της αριστερής εικόνας.

7.2.1 Υλοποίηση ανομοιότητας σε ΜΕΓ

Η σχεδίαση του αλγορίθμου ανομοιότητας στην πλατφόρμα CUDA ακολουθεί τις βελτιστοποιήσεις που περιγράφηκαν στην ενότητα 6.1, υιοθετώντας ξεχωριστούς πυρήνες επεξεργασίας ώστε να διευκολύνει την κατατομή του χρόνου εκτέλεσης.

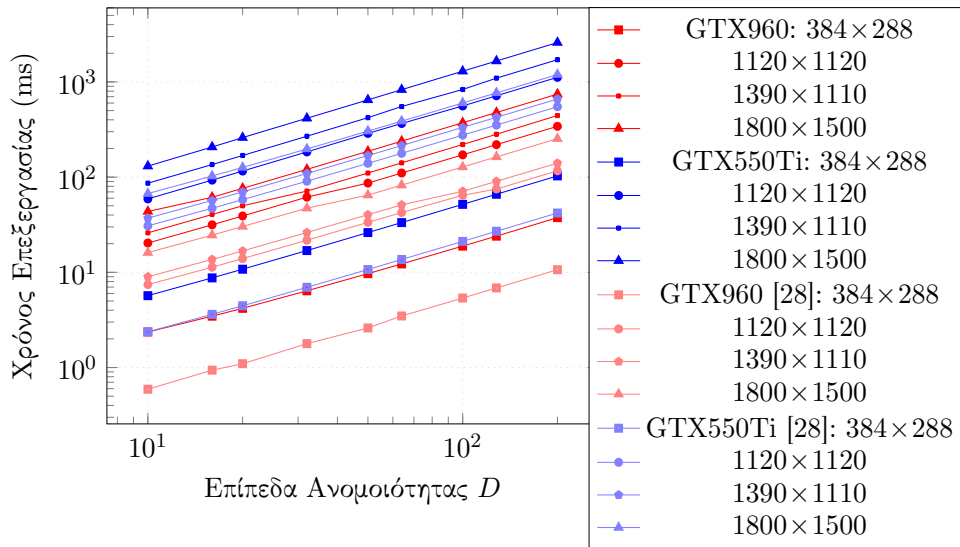
7.2. Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγόριθμους ΚΕΦΑΛΑΙΟ 7 σtereοσκοπικής αντιστοίχισης

Στην Εικ. 7.2 παρουσιάζεται η ροή των δεδομένων στους πυρήνες για την προτεινόμενη υλοποίηση της ανομοιότητας σε ΜΕΓ. Αρχικά το στερεοσκοπικό ζεύγος εικόνων αντιγράφεται στη μνήμη της ΜΕΓ από τη μνήμη κλειδωμένων σελίδων του συστήματος που φιλοξενεί τη ΜΕΓ. Μέσω του πυρήνα *buildRightCubeκ* υπολογίζεται η διαφορά μεταξύ δεξιάς και αριστερής εικόνας ($R-L$), παράγοντας ένα κύβο από D επίπεδα διαστάσεων $W \times H$ (D είναι το πλήθος των ανομοιοτήτων ανά εικονοστοιχείο). Στη συνέχεια ο πυρήνας *aggrCubeκ* πραγματοποιεί τρισδιάστατο φιλτράρισμα στον κύβο μέσω ενός μη διαχωρίσιμου φίλτρου 13×13 δειγμάτων. Εξαιτίας του σχετικά μεγάλου μεγέθους του φίλτρου, επιλέγουμε το σχήμα διαμοιρασμένης μνήμης των συγγραφέων του [25] και αντιγράφουμε το αποτέλεσμα σε ένα δεύτερο κύβο. Στην περίπτωση της *GTX 960*, τροποποιούμε την ανωτέρω μέθοδο ώστε να υπολογίζει 116×52 αποκρίσεις με κάθε νήμα να επεξεργάζεται 8 τιμές. Κατόπιν, ο πυρήνας *cubeMinκ* αναζητά την ελάχιστη τιμή ανομοιότητας ανά εικονοστοιχείο, επί του άξονα z του φιλτραρισμένου κύβου σε κάθε επίπεδο $W \times H$ εικονοστοιχείων (όσο και η απόσταση μεταξύ δύο διαδοχικών στοιχείων στον άξονα z κατά τη χρήση μονοδιάστατης μνήμης αποθήκευσης). Το αποτέλεσμα (θέση ελαχίστου) αποθηκεύεται σε μνήμη $W \times H$ θέσεων. Να σημειωθεί εδώ ότι εφόσον η τιμή του D αναμένεται τουλάχιστον δύο τάξεις μεγέθους μικρότερη από το πλήθος των εικονοστοιχείων σε κάθε επίπεδο, η εύρεση του ελαχίστου δεν απαιτεί τη χρήση παράλληλης περιστολής στον πυρήνα. Η ίδια διαδικασία (*Διαφορά, Διαστολή, Συνέλιξη and Αναζήτηση Ελαχίστου*) επαναλαμβάνεται για τον κύβο που αποθηκεύει τη διαφορά αριστερού και δεξιού ζεύγους ($L-R$). Με τη χρήση παρόμοιας τεχνικής με εκείνη της Εικ. 6.5, ο πυρήνας *crossCheckκ* συγκρίνει τις συντεταγμένες ελαχίστης ανομοιότητας όπως αυτές ανιχνεύθηκαν από τους πυρήνες *cubeMinκ*. Το δυαδικό αποτέλεσμα πολλαπλασιάζεται με την ελάχιστη ανομοιότητα του αριστερού κύβου έτσι ώστε όλες οι τιμές ανομοιότητας σε διαφορετικές συντεταγμένες να μηδενίζονται. Τέλος, ο πυρήνας *parPeakInterpκ* εκτελεί την παραβολική παρεμβολή των ανομοιοτήτων. Συνολικά, η περιγραφείσα υλοποίηση χρησιμοποιεί 2 κύβους μεγέθους $W \times H \times D$ -bytes καθώς μόνο η διαφορά $L-R$ επηρεάζει το τελικό αποτέλεσμα της παρεμβολής.

Διερευνήθηκε επίσης μια εναλλακτική υλοποίηση ανομοιότητας, στην οποία δεσμεύεται μνήμη για ένα μόνο κύβο. Σε αυτή την περίπτωση, το αποτέλεσμα τη συνέλιξης για κάθε επίπεδο του κύβου αποθηκεύεται σε μνήμη $W \times H$ δειγμάτων και αντιγράφεται στην αντίστοιχη θέση μνήμης εντός του κύβου. Αυτό έχει ως αποτέλεσμα την εξοικονόμηση $W \times H \times D \times 4$ bytes μνήμης στη ΜΕΓ με κόστος D αντιγραφές μεγέθους $W \times H \times 4$ bytes εσωτερικά της ΜΕΓ. Βάσει των πειραματικών μας μετρήσεων, σε όλους τους πυρήνες δισδιάστατης επεξεργασίας υιοθετούμε blocks 32×4 επεξεργαστικών νημάτων στις *GTX 670* και *GTX 960* και 32×8 στην *GTX 550Ti*. Οι πυρήνες τρισδιάστατης επεξεργασίας χρησιμοποιούν ένα νήμα στον άξονα z και D blocks. Οι πυρήνες φιλτραρίσματος βασίζονται σε blocks 32×32 επεξεργαστικών νημάτων.

7.2. Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγόριθμους
 ΚΕΦΑΛΑΙΟ 7 σtereοσκοπικής αντιστοίχισης

Επιδόσεις Ανομοιότητας σε MEG ($10 \leq D \leq 200$)



Εικόνα 7.3: Χρόνος εκτέλεσης υλοποιήσεων αλγορίθμου ανομοιότητας σε MEG.

Αποτίμηση Επιδόσεων

Για την αποτίμηση των επιδόσεων των υλοποιήσεων ανομοιότητας στην πλατφόρμα CUDA, καταμετρήθηκε ο χρόνος εκτέλεσης με τιμές ανομοιότητας $10 \leq D \leq 200$ στις προαναφερθείσες MEG, από την εκκίνηση του πυρήνα *buildRightCubeκ* έως τον πυρήνα *parPeakInterpκ* αγνοώντας τις αντιγραφές από και προς τη MEG. Χρησιμοποιήθηκαν 12 εικόνες από τη βάση στερεοσκοπικών εικόνων Middlebury [77, 78, 79], διαστάσεων 348×288 έως 1800×1500 . Πιο συγκεκριμένα, χρησιμοποιήθηκαν οι εικόνες *books*, *computer*, *cones*, *dolls*, *drumsticks*, *dwarves*, *laundry*, *moebius*, *reindeer*, *teddy*, *tsukuba* και *vintage*. Τα αποτελέσματα της υλοποίησης ενός κύβου διαστάσεων $W \times H \times D$ παρουσιάζονται στην Εικόνα 7.3. Στη χαμηλότερη ευκρίνεια των 384×288 εικονοστοιχείων, η προτεινόμενη υλοποίηση απαιτεί 3.4ms όταν $D=10$, 8.756ms όταν $D=32$ και 53.07ms όταν $D=200$ στη GTX 670. Σε αναλύσεις μέσης ευκρίνειας (π.χ. 1120×1120) ο χρόνος υπολογισμού κυμαίνεται από 30.6ms ($D=10$) έως 81.28ms ($D=32$) και φτάνει τα 498.4ms ($D=200$). Τα αποτελέσματα καταδεικνύουν γραμμική εξάρτηση του χρόνου υπολογισμού τόσο από την ανάλυση, όσο και από το πλήθος των τιμών ανομοιότητας. Για παράδειγμα σε εικόνες διαστάσεων 1800×1500 , η ίδια MEG απαιτεί από 67.16 ($D=10$) έως 1119.9ms ($D=200$). Η GTX 550Ti είναι κατά περίπου 1.66x-2x βραδύτερη σε όλες τις περιπτώσεις. Να σημειωθεί ότι λόγω των αυξημένων απαιτήσεων αποθήκευσης του αλγορίθμου ανομοιότητας η GTX 960 είναι κατά μέσο όρο κατά 1.4x ταχύτερη από τη GTX 670 καθώς διαθέτει μεγαλύτερη ενοποιημένη λανθάνουσα μνήμη μεγέθους 24KB, και περισσότερη διαμοιρασμένη μνήμη ανά νηματικό πολυεπεξεργαστή. Η βασισμένη σε δύο κύβους υλοποίηση είναι κατά περίπου 1.09x ταχύτερη σε όλες τις MEG.

Στο γράφημα της εικόνας 7.3 παρουσιάζονται επίσης τα αποτελέσματα της

Πίνακας 7.6: Ρυθμός πλαισίων ανομοιότητας σε ΜΕΓ.

ΜΕΓ	Σχετική Ισχύς ΜΕΓ ¹	Μέθοδος	Π.Α.Δ.
GTX 960	1	Προτεινόμενη	225
GTX 960	1	[28] (Box φίλτρο)	807
GTX 680	0.966	[50]	90
GTX 580	0.789	[49]	62
GTX 480	0.693	[36]	24

μεθόδου [28], επεξεργαστικά παραπλήσιας με την προτεινόμενη (υπολογίζει δύο χάρτες βάθους, υλοποιεί αντιστοίχιση με επεξεργασία παραθύρου και διαθέτει κλασματική ακρίβεια). Για την επίτευξη δίκαιων συγκρίσεων απόδοσης, ο κώδικας της μεθόδου [28] μεταγλωττίστηκε για τις προαναφερθείσες ΜΕΓ τροποποιώντας τις διαστάσεις του παραθύρου άθροισης σε 13×13 δείγματα. Από τις μετρήσεις μας προκύπτει ότι η μέθοδος στο [28] είναι έως και 2.9x ταχύτερη σε εικόνες διαστάσεων 1120×1120 και άνω και έως 3.9x ταχύτερη σε εικόνες 384×288 . Τα αποτελέσματα αυτά είναι αναμενόμενα, καθώς οι συγγραφείς της μεθόδου χρησιμοποιούν διαχωρίσιμο φίλτρο και υπολογίζουν το τρέχον άθροισμα αντί του προτεινόμενου πολ/μού ακεραίων. Η προτεινόμενη υλοποίηση επιτυγχάνει απόδοση της τάξης των 38.48 π.α.δ. σε εικόνες 1390×1110 δειγμάτων ($D = 10$) και υπερβαίνει τα 423 π.α.δ. για εικόνες 384×288 στο ίδιο επίπεδο ανομοιότητας.

Ο πίνακας 7.6 παρουσιάζει τις επιδόσεις της προτεινόμενης υλοποίησης του αλγορίθμου ανομοιότητας κατά την εκτέλεση στην GTX 960 εν συγκρίσει με υλοποιήσεις παραπλήσιων αλγορίθμων στη βιβλιογραφία. Όλες οι μέθοδοι περιλαμβάνουν άθροιση βασισμένη σε παράθυρο και οι [50, 36] συνδυάζουν επεξεργασία στο πεδίο του χώρου και του χρόνου. Επίσης, όλες οι μέθοδοι χρησιμοποιούν εικόνες διαστάσεων 320×240 και τιμή ανομοιότητας $D = 32$. Για διευκόλυνση των συγκρίσεων, παρουσιάζουμε επίσης ένα σχετικό δείκτη επιδόσεων των τεσσάρων ΜΕΓ του πίνακα 7.6. Οι σχετικές επιδόσεις των ΜΕΓ υπολογίστηκαν από την τιμή του διάμεσου θεωρώντας σημείο αναφοράς τη GTX 960 και κοινό πρόγραμμα συγκριτικής αξιολόγησης το οποίο εκτελεί τον αλγόριθμο οπτικής ροής TV-L1 Optical Flow (αποτελέσματα από τη σελίδα [47]). (Να σημειωθεί εδώ ότι ο δείκτης αυτός δεν αποτελεί απόλυτη ένδειξη των σχετικών επιδόσεων, καθώς οι ΜΕΓ του πίνακα 7.6 αντιπροσωπεύουν πολλαπλές γενεές αρχιτεκτονικών).

7.2.2 Υλοποίηση ανομοιότητας σε FPGA

Για την επιτάχυνση του αλγορίθμου ανομοιότητας σε FPGA αναπτύχθηκε μια σωληνωτή αρχιτεκτονική σε επίπεδο εικονοστοιχείου, βασισμένη σε α) παράλληλους αριθμητικούς υπολογισμούς και το βασικότερο, β) κατάτμηση δύο επιπέδων του κύβου ανομοιοτήτων με στόχο την ελαχιστοποίηση της απαιτούμενης μνήμης.

¹Η εκτίμηση πραγματοποιήθηκε βάσει της κατάταξης στη διεύθυνση <https://compubench.com/> και των επιδόσεων στον αλγόριθμο οπτικής ροής TV-L1 Optical Flow.

Αναλυτικότερα, εκμεταλλευόμαστε το γεγονός ότι η ανίχνευση της ελάχιστης μετρικής διά μέσου των D επιπέδων γίνεται με βάση το μοντέλο του νικητή και ότι οι εικόνες επιδιορθώνονται. Το μοντέλο του νικητή επιτρέπει τον επαναληπτικό υπολογισμό κάθε ενός εκ των D επιπέδων και την ενημέρωση μόνο των μετρικών που βελτιώνουν το τρέχον ελάχιστο. Συνεπώς, αποθηκεύονται μόνο $W \times H$ ολικά ελάχιστα αντί όλου του χώρου ανομοιότητας δηλαδή μόνο ένα επίπεδο αντί για D). Η επιδιόρθωση των εικόνων επιτρέπει την κατάτμηση του επιπέδου των εικονοστοιχείων σε s οριζόντιες λωρίδες διαστάσεων $W \times H/s$ και την επεξεργασία αυτών διαδοχικά και ανεξάρτητα. Καλούμε τον αναγνώστη να παρατηρήσει ότι οι διαδοχικές λωρίδες πρέπει να επικαλύπτονται κατά $M/2 - 1 = 3$ γραμμές ώστε το παράθυρο άθροισης $M \times M$ δειγμάτων να μετακινείται αδιάλειπτα δια μέσου των λωρίδων. Η περιγραφείσα κατάτμηση δύο επιπέδων του χώρου ανομοιότητας επιτρέπει την αποθήκευση $W \times H/s$ τιμών ανομοιότητας αντί $D \times W \times H$, δηλαδή βελτιστοποιεί τις απαιτήσεις μνήμης κατά 4 τάξεις μεγέθους.

Με βάση την προσέγγιση που μόλις περιγράφηκε, αναπτύχθηκε σωληνωτή αρχιτεκτονική η οποία επεξεργάζεται ένα εικονοστοιχείο ανά κύκλο από κάθε εικόνα με σειρά σάρωσης πλεγματοειδούς εικόνας (raster-scan order). Δηλαδή, παράγει μία τιμή μετρικής ομοιότητας ανά κύκλο και συνεπώς συντίθεται το i -οστό επίπεδο του κύβου, όπου $0 < i < D$ (επίσης με σειρά σάρωσης πλεγματοειδούς εικόνας). Η σωλήνωση χρειάζεται i διαδοχικές επαναλήψεις, μία για κάθε επίπεδο του κύβου. Σε κάθε επανάληψη, αναγιγνώσκονται τα εικονοστοιχεία της αριστερής εικόνας στη συντεταγμένη $x = 0$ και εκείνα της δεξιάς εικόνας στη συντεταγμένη $x = -i$, δημιουργώντας μία εικονική μετατόπιση κατά i εικονοστοιχεία. Κατά παρόμοιο τρόπο με τη συνέλιξη του αλγορίθμου *L-SEABI* η αρχιτεκτονική διατηρεί ρυθμοαπόδοση 1 μετρικής/κύκλο μέσω πλήρους παραλληλοποίησης του παραθύρου των 7×7 γκαουσιανών συντελεστών και ενός σωληνωτού δένδρου άθροιστών 7×7 σε 1. Της παράλληλης αυτής υπολογιστικής μονάδας προηγείται μία προσωρινή μνήμη αποσειριοποίησης από 1 σε 7×7 στην είσοδο της οποίας εισαγάγονται σειριακά οι διαφορές μεταξύ αριστερής και δεξιάς εικόνας. Εν συνεχεία, οι τιμές των μετρικών προωθούνται από το δένδρο άθροισης σε μία μονάδα *ενημέρωσης χάρτη βάθους*, η οποία αποθηκεύει μόνο την ελάχιστη τιμή μετρικής ανά εικονοστοιχείο και τη συγκρίνει με τους επόμενους υπολογισμούς για πιθανές ενημερώσεις. Επιπρόσθετα, η μονάδα *ενημέρωσης* αποθηκεύει την προηγούμενη τιμή ελαχίστου (η οποία αποθηκεύεται προσωρινά κατά την προηγούμενη επανάληψη) και υπολογίζει απευθείας την τιμή της παραβολικής παρεμβολής όταν εισαχθεί η τρίτη τιμή δείγματος (κατά την επόμενη επανάληψη). Κατά συνέπεια, ο απευθείας υπολογισμός της παραβολικής παρεμβολής διπλασιάζει τις απαιτήσεις σε μνήμη (εφόσον χρειάζεται δύο επίπεδα από το χώρο ανομοιότητας). Τα αποτελέσματα ανομοιότητας της αριστερής εικόνας αποθηκεύονται στην επιπλίνθια μνήμη και οι πόροι της σωλήνωσης επαναχρησιμοποιούνται για τον υπολογισμό του δεξιού χάρτη βάθους μετά από D νέες επαναλήψεις. Στο τελικό στάδιο πραγματοποιείται έλεγχος συνέπειας των δύο χαρτών. Μετά από τον υπολογισμό του χάρτη ανομοιότητας μίας λωρίδας $W \times H/s$, οι πόροι του υλικού επαναχρησιμοποιούνται για την επόμενη λωρίδα έως την επεξεργασία ολόκληρης της εικόνας. Για περισσότερες λεπτομέρειες επί της προτεινόμενης αρχιτεκτονικής ο αναγνώστης

μπορεί να ανατρέξει στη δημοσίευση [51].

Η προαναφερθείσα μονάδα ανομοιότητας υλοποιήθηκε με τη βοήθεια παραμετρικού κώδικα περιγραφής υλικού στο FPGA ολοκληρωμένο Xilinx Virtex xc6v1x240t-2. Το κόστος υλοποίησης για εικόνες μεγέθους 1120×1120 , $D = 200$ και διαστάσεις λωρίδας 1120×28 , ανέρχεται σε 998 slices (3% επί του συνολικού), 2978 LUTs (2%), 3116 καταχωρητές (1%), 0 μονάδες DSPs, και 101 μονάδες μνήμης RAMB36 (24%). Η μεγέθυνση του παραθύρου άθροισης από 7×7 σε 13×13 δείγματα, αυξάνει το κόστος του υλικού σε 2690 slices (8,492 LUTs και 8,640 καταχωρητές) από όπου προκύπτει ότι το πιο δαπανηρό τμήμα της αρχιτεκτονικής αποτελεί η μονάδα παράλληλων αριθμητικών υπολογισμών υπεύθυνη για την άθροιση (αναλόγως του μεγέθους του παραθύρου, η άθροιση απαιτεί περίπου το 60%-85% από τους συνολικούς αναλισκόμενους πόρους). Η μέγιστη επιτευχθείσα συχνότητα λειτουργίας ανέρχεται στα 344 MHz, επιτρέποντας στην προτεινόμενη μονάδα ανομοιότητας να επεξεργαστεί πλήρως το ζεύγος εικόνων μεγέθους 1120×1120 σε 1.87 δευτερόλεπτα (ο χρόνος είναι σχεδόν ανεξάρτητος του μεγέθους του παραθύρου εξαιτίας της πλήρους παραλληλοποίησης αυτού).

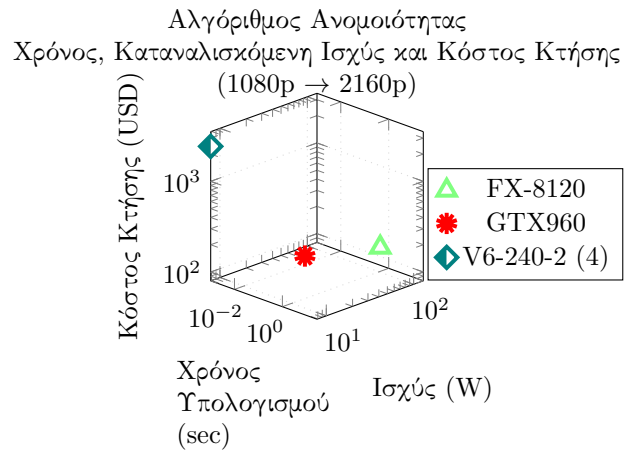
Περαιτέρω παραλληλοποίηση της αρχιτεκτονικής σε επίπεδο εικόνας, με τη χρήση πολλαπλών μονάδων ανομοιότητας να επεξεργάζονται παράλληλα πολλαπλές λωρίδες από την εικόνα, ελαττώνει το χρόνο επεξεργασίας σε 0.54 sec, με κατανάλωση περίπου 8.8 Watts, για $P = 4$ μονάδες ανομοιότητας στο FPGA xc6v1x240t-2. Η προτεινόμενη αρχιτεκτονική αποδεικνύεται αποδοτική σε αναλισκόμενους πόρους, καθώς απαιτεί 20x λιγότερες λογικές μονάδες LUT, 0 μονάδες DSP και το 1/3 των μονάδων μνήμης RAMB σε σύγκριση με την αρχιτεκτονική στο [42] (η οποία ωστόσο χρειάζεται μόλις 4 ms για 64 επίπεδα ανομοιότητας). Επίσης, η προτεινόμενη αρχιτεκτονική απαιτεί 24x λιγότερες λογικές μονάδες LUT και 0 μονάδες DSP αντί των 625 της αρχιτεκτονικής στο [94] (η οποία επεξεργάζεται 1080p εικόνες με ρυθμό 30 π.α.δ.).

Αποτελέσματα υλοποίησης αλγορίθμου ανομοιότητας: Επιβεβαίωση των προτεινόμενων τεχνικών επιτάχυνσης

Όσον αφορά τα αποτελέσματα επιτάχυνσης του αλγορίθμου ανομοιότητας που παρουσιάζεται στην παρούσα ενότητα, η ελάχιστη καταμετρούμενη επιτάχυνση της MEG (*GTX 960* έναντι του επεξεργαστή *i5-3470*) για $D = 200$ μπορεί να φτάσει και το 50.33x σε εικόνες διαστάσεων 1800×1500 και το 45.5x σε εικόνες 1120×1120 για τις οποίες το FPGA xc6v1x240t-2 επιτυγχάνει επιτάχυνση 28.8x με χρήση 4 παράλληλων μονάδων.

Η αποτίμηση των αποτελεσμάτων υλοποίησης του αλγορίθμου ανομοιότητας επιβεβαιώνει τα αποτελέσματα του κεφαλαίου 6 όπως φαίνεται και στο ένθετο γράφημα της εικόνας 7.4 για εικόνες μεγέθους 1120×1120 δειγμάτων. Σε αυτή την περίπτωση η MEG *GTX 960* καταναλώνει περίπου 80 W και συνεπώς εμφανίζει απόδοση 9.17 MDEs/W (εκατομμύρια εκτιμήσεις ανομοιότητας ανά Watt, millions of disparity estimations per Watt) και παρόλο που το FPGA xc6v1x240t-2 εί-

7.2. Εφαρμογή των προτεινόμενων τεχνικών επιτάχυνσης σε αλγόριθμους
 ΚΕΦΑΛΑΙΟ 7 στερεοσκοπικής αντιστοίχισης



Εικόνα 7.4: Χρόνος και Ισχύς κατά τη διεργασία ΥΑ μέσω του αλγορίθμου *SIL-SEABI*. Στο γράφημα περιλαμβάνονται διάφορα ολοκληρωμένα κυκλώματα μαζί με το αντίστοιχο κόστος κτήσης τους (σε USD) (ο συμβολισμός εξηγείται στο κυρίως κείμενο).

ναι κατά περίπου 1.6x βραδύτερο, καταναλώνει μόλις 8.8 W δηλαδή επιτυγχάνει απόδοση της τάξης των 51.8 MDEs/W. Στον αντίποδα, η αντίστοιχη απόδοση του επεξεργαστή *FX-8120* είναι μόλις 0.11 MDEs/W.

Συμπεράσματα

Έχοντας ως στόχο την επεξεργασία εικόνων και εικονορροών υπό το πρίσμα της μειωμένης πολυπλοκότητας, στην παρούσα διατριβή προτάθηκαν αρχικά τρεις νέες τεχνικές παρεμβολής. Οι τεχνικές χρησιμοποιήθηκαν κατά την εκτίμηση κίνησης του προτύπου H.264/AVC, στα πλαίσια του οποίου αποτιμήθηκε η απόδοση και ο χρόνος εκτέλεσής τους. Τα αποτελέσματα συγκρίθηκαν με τρεις γνωστές τεχνικές παρεμβολής και οδήγησαν στο συμπέρασμα ότι τα φίλτρα δικυβικής παρεμβολής 4 δειγμάτων αποτελούν την επικρατέστερη μέθοδο αντικατάστασης των φίλτρων 6 δειγμάτων του προτύπου. Με τη εφαρμογή των φίλτρων αυτών στις προτεινόμενες ακμοπροσαρμοστικές τεχνικές επιτεύχθηκε περαιτέρω ελάττωση του χρόνου εκτίμησης κίνησης.

Εν συνεχεία, οι τεχνικές αυτές χρησιμοποιήθηκαν στο στάδιο υπερδειγματοληψίας σφάλματος του πλαισίου επαναληπτικής οπισθοπροβολής. Ως συνέπεια των ανωτέρω, ενισχύθηκε η ποιότητα της ανακατασκευασμένης εικόνας υψηλής ευκρίνειας για διάφορους παράγοντες υπερδειγματοληψίας. Σε σύγκριση με το συμβατικό πλαίσιο επαναληπτικής οπισθοπροβολής, επιτεύχθηκε μέση βελτίωση της ποιότητας κατά 2.2 dB με παράγοντα υπερδειγματοληψίας 4. Επιπρόσθετα, μειώθηκε ο συνολικός χρόνος ανακατασκευής με συνέπεια την αύξηση της συνολικής αποτελεσματικότητας.

Κατόπιν, εισήχθη ένας νέος αλγόριθμος υπερ-ανάλυσης, ο οποίος εξετάστηκε αρχικά ως αυτόνομη λύση υπερδειγματοληψίας εικόνων. Ο αλγόριθμος παρουσίασε ανταγωνιστικές επιδόσεις σε σχέση με τις μεθόδους τεχνολογικής αιχμής, τόσο όσον αφορά την ποιότητα, όσο και την πολυπλοκότητα, όπως η τελευταία αποτιμήθηκε μέσω του χρόνου εκτέλεσης. Ο αλγόριθμος σχεδιάστηκε και χρησιμοποιήθηκε προκειμένου να εμπλουτίσει το γενικευμένο σχήμα κωδικοποίησης με αποδεκάτιση, μέσω της αύξησης του κρίσιμου δυφορρυθμού και την ελάττωση του συνολικού χρόνου κωδικοαποκωδίκευσης έως και κατά 71%. Τα ανωτέρω αποτελέσματα οδήγησαν στη μελέτη και την πρόταση ενός αποτελεσματικού σχήματος συμπίεσης εικονορροών, το οποίο συνδυάζει υπερσύγχρονες μεθόδους κωδικοαποκωδίκευσης με το πεδίο της υπερ-ανάλυσης, έχοντας ως αποτέλεσμα τη βελτίωση της αποδοτικότητας κωδικοποίησης και της πολυπλοκότητας των συμβατικών προτύπων, σε εφαρμογές με απαιτήσεις χαμηλού δυφορρυθμού. Το προτεινόμενο σχήμα περιλαμβάνει την κωδικοποίηση εικόνων χαμηλότερης ανάλυσης και την υπερδειγματοληψία τους στην πλευρά του αποκωδικοποιητή. Η

ΚΕΦΑΛΑΙΟ 8

αρχική αποτίμηση αυτού του σχήματος με ευρέως διαδεδομένες τεχνικές παρεμβολής, κατέδειξε ότι μπορεί να βελτιώσει την ποιότητα κατά 2 έως 4 dB για δυοφορρυθμούς κάτω των 3 Mb/s, υποτετραπλασιάζοντας παράλληλα την πολυπλοκότητα της κωδικοποίησης.

Ο προτεινόμενος αλγόριθμος υπερ-ανάλυσης αναπτύχθηκε σε συνδυασμό με την εξερεύνηση/βελτιστοποίηση της διαδικασίας αποδεκτίσης, προκειμένου να παράσχει βελτιωμένη ποιότητα, να διατηρήσει την πολυπλοκότητα σε χαμηλά επίπεδα και να εμπλουτίσει το γενικευμένο σχήμα συμπίεσης. Με την προτεινόμενη τεχνική υπερ-ανάλυσης επιτυγχάνονται τα εξής:

- μέση βελτίωση της ποιότητας κατά 5 dB σε σχέση με απλές τεχνικές φιλτραρίσματος και βελτιωμένη ποιότητα σε σχέση με σχήματα επαναληπτικής οπισθοπροβολής,
- συγκρίσιμες τιμές PSNR σε σχέση με τις εξετασθείσες υπερσύγχρονες μεθόδους υπερ-ανάλυσης,
- παραπλήσιες τιμές MSSIM σε σχέση με τις εξετασθείσες υπερσύγχρονες μεθόδους υπερ-ανάλυσης από μία εικόνα,
- τα καλύτερα αποτελέσματα ποιότητας βάσει της μετρικής BRISQUE σε σχέση με τις εξετασθείσες υπερσύγχρονες μεθόδους υπερ-ανάλυσης από μία εικόνα,
- ελάττωση του χρόνου εκτέλεσης κατά μία έως τρεις τάξεις μεγέθους σε σχέση με τις υπερσύγχρονες μεθόδους, λόγω της χρήσης αριθμητικής ακεραίων και μετασχηματισμών χαμηλής πολυπλοκότητας.

Η ενσωμάτωση του προτεινόμενου αλγορίθμου στο βελτιωμένο σχήμα συμπίεσης, οδήγησε σε ένα υβριδικό σχήμα το οποίο υπερσχύει των προτύπων H.264 και HEVC για δυοφορρυθμούς κάτω των 10 και 3.8 Mb/s αντιστοίχως. Όλες οι ανωτέρω βελτιώσεις καταδεικνύουν ότι το προτεινόμενο σχήμα αποτελεί μία πολλά υποσχόμενη λύση για εφαρμογές μετάδοσης εικονορροών περιορισμένου εύρους ζώνης και κωδικοποιητές χαμηλής πολυπλοκότητας. Εκτός των βελτιώσεων εντός του σχήματος συμπίεσης, η προτεινόμενη τεχνική απεδείχθη αποτελεσματική ως αυτόνομη λύση υπερ-ανάλυσης η οποία δύναται επίσης να χρησιμοποιηθεί και για τη βελτίωση της απόδοσης των μεθόδων υπερ-ανάλυσης τεχνολογικής αιχμής.

Εν συνεχεία η διατριβή μελέτησε τα απτά αποτελέσματα των ανωτέρω σε πλατφόρμες επιτάχυνσης υλικού. Με βάση τα αποτελέσματα της μεθόδου *L-SEABI*, προτάθηκαν τεχνικές παραλληλοποίησης και βελτιστοποιήσεις αυτών για εφαρμογές σε μονάδες επεξεργασίας γραφικών και συστοιχίες επιτόπια προγραμματιζόμενων πυλών (FPGA). Οι προτεινόμενες τεχνικές επιτάχυνσης σε μονάδες επεξεργασίας γραφικών αποδείχθηκαν μεθοδολογικά έγκυρες για ευρύ πεδίο αρχιτεκτονικών, συμπεριλαμβανομένων και των πλέον σύγχρονων. Εστιάζοντας σε

πολλαπλά επίπεδα, από εκείνο της αλγοριθμικής σχεδίασης έως τη διεπαφή προγραμματισμού εφαρμογής, οι προτεινόμενες τεχνικές σχεδιάστηκαν με στόχο την αξιοποίηση των αρχιτεκτονικών των μονάδων επεξεργασίας γραφικών. Συνδυάζουν βελτιωμένη ρυθμοαπόδοση, παραλληλισμό επιπέδου εντολής, με μειωμένο λανθάνοντα χρόνο και μειωμένη νηματική απόκλιση. Σε σύγκριση με τη συμβατική απαίτηση επεξεργασίας πραγματικού χρόνου των 30 π.α.δ., οι προτεινόμενες τεχνικές επιτυγχάνουν ανακατασκευή εικόνων 3840×2160 με απόδοση της τάξης των 109 π.α.δ. σε μέσης απόδοσης και παλαιότερες μονάδες και 345 π.α.δ. σε υψηλότερης απόδοσης μονάδες επεξεργασίας γραφικών. Η προτεινόμενη παραμετροποιήσιμη και κλιμακοθετήσιμη αρχιτεκτονική σε FPGA, εξετάστηκε για διάφορους βαθμούς παραλληλίας και σε διάφορα ολοκληρωμένα FPGA. Με τη βελτιστοποιημένη σωληναγωγή σε επίπεδο εικονοστοιχείου και διεργασίας, η συμβατική απόδοση πραγματικού χρόνου τετραπλασιάστηκε σε ολοκληρωμένα παλαιότερης γενιάς / χαμηλής κατανάλωσης και επετεύχθη έως και 69 φορές ταχύτερη επεξεργασία στο ολοκληρωμένο Virtex 7 2000t. Επιπλέον, παρουσιάστηκαν αποτελέσματα συγκριτικής αποτίμησης μεταξύ υλοποιήσεων σε επεξεργαστές γενικού σκοπού, μονάδες επεξεργασίας γραφικών και FPGA, ως προς την καταναλισκόμενη ισχύ κάθε πλατφόρμας. Συνεπώς, δόθηκε και μία εποπτική εικόνα της σχετικής αποτελεσματικότητας ισχύος κάθε πλατφόρμας.

Οι προτεινόμενες τεχνικές υπερ-ανάλυσης μπορούν να χρησιμοποιηθούν για την αντικατάσταση του αρχικού σταδίου ανακατασκευής των μεθόδων τεχνολογικής αιχμής. Όπως έδειξε η σύντομη μελέτη στο κεφάλαιο 7, οι τεχνικές διατηρούν τη σχετική απόδοσή τους και κατά τη σύγκριση με πρόσφατα προταθείσες μεθόδους υπερ-ανάλυσης τεχνολογικής αιχμής. Επιπλέον, η αντικατάσταση της δικυβικής παρεμβολής με τον αλγόριθμο *L-SEABI*, οδηγεί σε βελτίωση της ποιότητας όπως αυτή μετράται με αντικειμενικές μετρικές. Περαιτέρω διερεύνηση της βελτίωσης των μεθόδων τεχνολογικής αιχμής αποτελεί μελλοντικό ερευνητικό στόχο.

Τέλος, τα αποτελέσματα των στρατηγικών παραλληλοποίησης επιβεβαιώθηκαν με την εφαρμογή τους στον αλγόριθμο ανομοιότητας για τον υπολογισμό χάρτη βάθους βασιζόμενο σε μη διαχωρίσιμη συνέλιξη, έλεγχο συνέπειας από αριστερή και δεξιά εικόνα και ακρίβεια ενός τετάρτου του εικονοστοιχείου. Με την προτεινόμενη υλοποίηση σε μονάδες επεξεργασίας γραφικών, επετεύχθη επιτάχυνση κατά τουλάχιστον 14 φορές σε σχέση με τον ταχύτερο επεξεργαστή γενικού σκοπού (i5-3470) στη μέσης απόδοσης μονάδα GTX 550 Ti για εικόνες διαστάσεων 1120×1120 και 200 επίπεδα ανομοιότητας. Η προτεινόμενη αρχιτεκτονική σε FPGA οδήγησε σε αντίστοιχη επιτάχυνση κατά 29 φορές κατά την υλοποίηση σε ολοκληρωμένο μέσης απόδοσης. Μελλοντικά, οι προτεινόμενες τεχνικές παραλληλοποίησης θα εφαρμοστούν και θα αποτιμηθούν σε αλγορίθμους τηλεπικοινωνιακών συστημάτων ασύρματων και οπτικών επικοινωνιών.

ΚΕΦΑΛΑΙΟ 8

Βιβλιογραφία

- [1] H.264/AVC JM Reference Software, Aug. 2008.
- [2] KTA Software, 2008.
- [3] Advanced video coding for generic audiovisual services. 167–169.
- [4] NVIDIA’s Next Generation CUDA Compute Architecture: Kepler GK110, Aug. 2012.
- [5] H.265/HEVC HM Reference Software, May 2013.
- [6] ALEX, K. CUDA Convolutional Neural Networks, May 2015.
- [7] ANGELOPOULOU, M. E., BOUGANIS, C.-S., CHEUNG, P. Y. K., AND CONSTANTINIDES, G. A. Robust real-time super-resolution on fpga and an application to video enhancement. *ACM Trans. Reconfigurable Technol. Syst.* 2, 4 (Sept. 2009), 22–29.
- [8] ATES, H. F. Enhanced low bitrate h.264 video coding using decoder-side super-resolution and frame interpolation. *Optical Engineering* 52, 7 (2013), 071505–071505.
- [9] AUDIOVISUAL COMMUNICATIONS LABORATORY, E. P. F. D. L. Super-Resolution v2.0 Software.
- [10] BABACAN, S., MOLINA, R., AND KATSAGGELOS, A. Variational bayesian super resolution. *Image Processing, IEEE Transactions on* 20, 4 (2011), 984–999.
- [11] BARRETO, D., ALVAREZ, L. D., MOLINA, R., KATSAGGELOS, A. K., AND CALLICO, G. M. Region-based super-resolution for compression. *Multidimensional Systems and Signal Processing, Springer* 18, 2-3 (2007), 59–81.

- [12] BOWEN, O., AND BOUGANIS, C. Real-time image super resolution using an fpga. In *Field Programmable Logic and Applications, 2008. FPL 2008. International Conference on* (Sept 2008), pp. 89–94.
- [13] BRUCKSTEIN, A. M., ELAD, M., AND KIMMEL, R. Down-scaling for better transform compression. *IEEE Transactions on Image Processing* 12, 9 (September 2003), 1132–1144.
- [14] BURGER, W., AND BURGE, M. *Digital Image Processing, an Algorithmic approach using Java*, first ed. Springer, 2008.
- [15] CHE, S., LI, J., SHEAFFER, J., SKADRON, K., AND LACH, J. Accelerating compute-intensive applications with gpus and fpgas. In *Application Specific Processors, 2008. SASP 2008. Symposium on* (June 2008), pp. 101–107.
- [16] CHEN, T.-C., HUANG, Y.-W., AND CHEN, L.-G. Fully utilized and reusable architecture for fractional motion estimation of h.264/avc. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on* (May 2004), vol. 5, pp. V–9–12 vol.5.
- [17] DAI, S., HAN, M., WU, Y., AND GONG, Y. Bilateral back-projection for single image super resolution. In *Multimedia and Expo, 2007 IEEE International Conference on* (July 2007), pp. 1039–1042.
- [18] DODGSON, N. Quadratic interpolation for image resampling. *Image Processing, IEEE Transactions on* 6, 9 (Sep 1997), 1322–1326.
- [19] DONG, C., LOY, C., HE, K., AND TANG, X. Learning a deep convolutional network for image super-resolution. In *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., vol. 8692 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014, pp. 184–199.
- [20] DONG, C., LOY, C., HE, K., AND TANG, X. Image super-resolution using deep convolutional networks. *Pattern Analysis and Machine Intelligence, IEEE Transactions on PP*, 99 (2015), 1–1.
- [21] DONG, J., AND YE, Y. Adaptive downsampling for high-definition video coding. *Circuits and Systems for Video Technology, IEEE Transactions on* 24, 3 (March 2014), 480–488.
- [22] DONG, W., ZHANG, D., SHI, G., AND WU, X. Nonlocal back-projection for adaptive image enlargement. In *Image Processing (ICIP), 2009 16th IEEE International Conference on* (2009), pp. 349–352.
- [23] DONG, W., ZHANG, D., SHI, G., AND WU, X. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive

- regularization. *Image Processing, IEEE Transactions on* 20, 7 (2011), 1838–1857.
- [24] DONG, W., ZHANG, L., LUKAC, R., AND SHI, G. Sparse representation based image interpolation with nonlocal autoregressive modeling. *Image Processing, IEEE Transactions on* 22, 4 (April 2013), 1382–1394.
- [25] EKLUND, A., AND DUFORT, P. *GPU-Pro 5: Advanced Rendering Techniques - Non-Separable 2D, 3D and 4D filtering with CUDA*, first ed. CRC Press, 2014, ch. 5.
- [26] ELAD, M., AND FEUER, A. Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images. *IEEE Trans. on Image Processing* 6, 2 (December 1997), 1646–1658.
- [27] FARSIU, S., ROBINSON, D., ELAD, M., AND MILANFAR, P. Fast and robust multi-frame super-resolution. *IEEE Trans. on Image Processing* 13, 10 (October 2004), 1327–1344.
- [28] GALLUP, D., FRAHM, J.-M., AND STAM, J. Cuda stereo. In *nVidia GPU Technology Conference 2009* (September 2009).
- [29] GEORGIS, G., LENTARIS, G., AND REISIS, D. Study of interpolation filters for motion estimation with application in h.264/avc encoders. In *Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on* (2011), pp. 9–12.
- [30] GEORGIS, G., LENTARIS, G., AND REISIS, D. Single-image super-resolution using low-complexity adaptive iterative back-projection. In *Digital Signal Processing. 18th IEEE International Conference on* (2013), pp. 152–157.
- [31] GEORGIS, G., LENTARIS, G., AND REISIS, D. Reduced complexity superresolution for low-bitrate video compression. *IEEE Transactions on Circuits and Systems for Video Technology* 26, 2 (Feb 2016), 332–345.
- [32] GREISEN, P., HEINZLE, S., GROSS, M., AND BURG, A. An fpga-based processing pipeline for high-definition stereo video. *EURASIP Journal on Image and Video Processing* 2011, 1 (2011).
- [33] GUPTA, P. S., AND KORADA, R. Novel algorithm to reduce the complexity of quarter-pixel motion estimation, 2004.
- [34] GURUMANI, S., CHOLAKKAL, H., LIANG, Y., RUPNOW, K., AND CHEN, D. High-level synthesis of multiple dependent cuda kernels on fpga. In *Design Automation Conference (ASP-DAC), 2013 18th Asia and South Pacific* (Jan 2013), pp. 305–312.
- [35] HARRIS, M. Optimizing Parallel Reduction in CUDA, Mar. 2007.

- [36] HOSNI, A., RHEMANN, C., BLEYER, M., AND GELAUTZ, M. Temporally consistent disparity and optical flow via efficient spatio-temporal filtering. In *Advances in Image and Video Technology*, Y.-S. Ho, Ed., vol. 7087 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 165–177.
- [37] HUANG, Y.-W., HSIEH, B.-Y., CHIEN, S.-Y., MA, S.-Y., AND CHEN, L.-G. Analysis and complexity reduction of multiple reference frames motion estimation in h.264/avc. *Circuits and Systems for Video Technology, IEEE Transactions on* 16, 4 (April 2006), 507–522.
- [38] HYUN, C., AND SUNWOO, M. Low power complexity-reduced me and interpolation algorithms for h.264/avc. *Journal of Signal Processing Systems* 56, 2-3 (2009), 285–293.
- [39] ILGIN, H. A., AND CHAPARRO, L. F. Low bit rate video coding using dct-based fast decimation/interpolation and embedded zerotree coding. *IEEE Transactions on Circuits and Systems for video technology* 17, 7 (July 2007), 833–844.
- [40] IRANI, M., AND PELEG, S. Improving resolution by image registration. *CVGIP: Graph. Models Image Process.* 53, 3 (Apr. 1991), 231–239.
- [41] ITU-TELECOMMUNICATIONS STANDARDIZATION SECTOR STUDY GROUP 16 QUESTION 6, V. C. E. G. V. Motion compensated prediction with 1/8-pel displacement vector resolution, Oct 2006.
- [42] JIN, S., CHO, J., PHAM, X. D., LEE, K. M., PARK, S.-K., KIM, M., AND JEON, J. W. Fpga design and implementation of a real-time stereo vision system. *Circuits and Systems for Video Technology, IEEE Transactions on* 20, 1 (2010), 15–26.
- [43] JONES, D., POWELL, A., BOUGANIS, C., AND CHEUNG, P. Gpu versus fpga for high productivity computing. In *Field Programmable Logic and Applications (FPL), 2010 International Conference on* (Aug 2010), pp. 119–124.
- [44] KALAROT, R., AND MORRIS, J. Comparison of fpga and gpu implementations of real-time stereo vision. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on* (June 2010), pp. 9–15.
- [45] KAO, C.-Y., WU, C.-L., AND LIN, Y.-L. A high-performance three-engine architecture for h.264/avc fractional motion estimation. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on* 18, 4 (April 2010), 662–666.
- [46] KEYS, R. Cubic convolution interpolation for digital image processing. *Acoustics, Speech and Signal Processing, IEEE Transactions on* 29, 6 (Dec 1981), 1153–1160.

- [47] KISHONTI LTD. Compubench, a professional OpenCL and Renderscript benchmark, Mar. 2015.
- [48] KLEPKO, R., WANG, D., AND HUCHET, G. Combining distributed video coding with super-resolution to achieve h.264/avc performance. *Journal of Electronic Imaging* 21, 1 (2012), 013011–1–013011–11.
- [49] KOWALCZUK, J., PSOTA, E., AND PEREZ, L. Real-time stereo matching on cuda using an iterative refinement method for adaptive support-weight correspondences. *Circuits and Systems for Video Technology, IEEE Transactions on* 23, 1 (Jan 2013), 94–104.
- [50] KOWALCZUK, J., PSOTA, E., AND PEREZ, L. Real-time temporal stereo matching using iterative adaptive support weights. In *Electro/Information Technology (EIT), 2013 IEEE International Conference on* (May 2013), pp. 1–6.
- [51] LENTARIS, G., DIAMANTOPOULOS, D., SIOZIOS, K., SOUDRIS, D., AND RODRIGÁLVAREZ, M. A. Hardware implementation of stereo correspondence algorithm for the exomars mission. In *Field Programmable Logic and Applications (FPL), 2012 22nd International Conference on* (2012), IEEE, pp. 667–670.
- [52] LEVON, J. Oprofile 1.0, a statistical profiler for Linux systems., Mar. 2015.
- [53] LIN, W., AND DONG, L. Adaptive downsampling to improve image compression at low bit rates. *IEEE Transactions on Image Processing* 15, 9 (September 2006), 2513–2521.
- [54] MA, S., ZHANG, L., ZHANG, X., AND GAO, W. Block adaptive super resolution video coding. *Advances in Multimedia Information Processing, Springer* 5879 (June 2009), 1048–1057.
- [55] MAHMOUDZADEH, A., AND KASHOU, N. Evaluation of interpolation effects on upsampling and accuracy of cost functions-based optimized automatic image registration. *International Journal of Biomedical Imaging* 2013, 395915 (2013), 96–114.
- [56] MATIC, T., ALEKSI, I., AND HOCENSKI, U. Cpu, gpu and fpga implementations of mald: Ceramic tile surface defects detection algorithm. *Automatika* 55, 1 (2014).
- [57] MITTAL, A., MOORTHY, A., AND BOVIK, A. No-reference image quality assessment in the spatial domain. *Image Processing, IEEE Transactions on* 21, 12 (Dec 2012), 4695–4708.
- [58] MOLINA, R., KATSAGGELOS, A. K., ALVAREZ, L. D., AND MATEOS, J. Toward a new video compression scheme using super-resolution. In

- Proceeding of SPIE, Visual Communications and Image Processing, vol. 6077* (2006).
- [59] NGUYEN, V.-A., TAN, Y.-P., AND LIN, W. Adaptive downsampling/upsampling for better video compression at low bit rate. In *IEEE International Symposium on Circuits and Systems* (2008), pp. 1624–1627.
- [60] NICKOLLS, J., BUCK, I., GARLAND, M., AND SKADRON, K. Scalable Parallel Programming with CUDA. *ACM Queue Magazine* 6, 2 (April 2008), 40–53.
- [61] NVIDIA. CUDA C Programming Guide, Mar. 2015.
- [62] NVIDIA. NVIDIA CUDA Fast Fourier Transform library (cuFFT), May 2015.
- [63] NVIDIA. Parallel Thread Execution ISA, Mar. 2015.
- [64] OHM, J. R., AND SULLIVAN, G. *MPEG video compression advances, The MPEG representation of digital media (chapter 3)*, first ed. Springer, 2012.
- [65] OKUHATA, H., IMAI, R., ISE, M., OMAKI, R. Y., NAKAMURA, H., HARA, S., AND SHIRAKAWA, I. Implementation of dynamic-range enhancement and super-resolution algorithms for medical image processing. In *Consumer Electronics (ICCE), 2014 IEEE International Conference on* (2014), IEEE, pp. 181–184.
- [66] PAN, Z., AND XIONG, H. Sparse spatio-temporal representation with adaptive regularized dictionaries for super-resolution based video coding. In *Data Compression Conference (DCC), 2012* (April 2012), pp. 139–148.
- [67] PARK, S. C., PARK, M. K., AND KANG, M. G. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE* 20, 3 (2003), 21–36.
- [68] PEREIRA, F., TORRES, L., GUILLEMOT, C., EBRAHIMI, T., LEONARDIE, R., AND KLOMP, S. Distributed video coding: Selecting the most promising application scenarios. *Signal Processing: Image Communication, Elsevier* 23, 5 (September 2008), 339–352.
- [69] PÉREZ, J., MAGDALENO, E., PÉREZ, F., RODRÍGUEZ, M., HERNÁNDEZ, D., AND CORRALES, J. Super-resolution in plenoptic cameras using fpgas. *Sensors* 14, 5 (2014), 8669–8685.
- [70] PIETRON, M., WIELGOSZ, M., ZUREK, D., JAMRO, E., AND WIATR, K. Comparison of gpu and fpga implementation of svm algorithm for fast image segmentation. In *Architecture of Computing Systems* ☒ *ARCS*

- 2013, vol. 7767 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 292–302.
- [71] PODLOZHNYUK, V. Image Convolution with CUDA, July 2012.
- [72] PROTTER, M., AND ELAD, M. Super resolution with probabilistic motion estimation. *IEEE Trans. on Image Processing* 18, 8 (August 2009), 1899–1904.
- [73] QIN, F.-Q., HE, X.-H., CHEN, W.-L., YANG, X.-M., AND WU, W. Video super-resolution reconstruction based on sub-pixel registration and iterative back projection. *Journal of Electronic Imaging* 18, 1 (2009), 013007–013007–11.
- [74] SANADA, Y., OHIRA, T., CHIKUDA, S., IGARASHI, M., IKEBE, M., ASAI, T., AND MOTOMURA, M. Fpga implementation of single-image super-resolution based on frame-bufferless box filtering. *Journal of Signal Processing* 17, 4 (2013), 111–114.
- [75] SANDERS, J., AND KANDROT, E. *CUDA by Example: An Introduction to General-Purpose GPU Programming*, 1st ed. Addison-Wesley Professional, 2010.
- [76] SAVARIMUTHU, T., KJ?R-NIELSEN, A., AND SORENSEN, A. Real-time medical video processing, enabled by hardware accelerated correlations. *Journal of Real-Time Image Processing* 6, 3 (2011), 187–197.
- [77] SCHARSTEIN, D., AND PAL, C. Learning conditional random fields for stereo. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on* (June 2007), pp. 1–8.
- [78] SCHARSTEIN, D., AND SZELISKI, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47, 1-3 (2002), 7–42.
- [79] SCHARSTEIN, D., AND SZELISKI, R. High-accuracy stereo depth maps using structured light. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on* (June 2003), vol. 1, pp. I–195–I–202 vol.1.
- [80] SCHWARZ, S., OLSSON, R., SJOSTROM, M., AND TOURANCHEAU, S. Adaptive depth filtering for hevc 3d video coding. In *Picture Coding Symposium (PCS), 2012* (2012), pp. 49–52.
- [81] SHEN, M., XUE, P., AND WANG, C. Down-sampling based video coding using super-resolution technique. *IEEE Transactions on Circuits and Systems for video technology* 21, 6 (June 2011), 755–765.

- [82] SU, D., AND WILLIS, P. Image interpolation by pixel-level data-dependent triangulation. *Computer Graphics Forum* 23, 2 (July 2004), 189–201. ID number: ISI:000222579000006.
- [83] SULLIVAN, G., OHM, J., HAN, W.-J., AND WIEGAND, T. Overview of the high efficiency video coding (hevc) standard. *Circuits and Systems for Video Technology, IEEE Transactions on* 22, 12 (2012), 1649–1668.
- [84] SZELISKI, R. *Computer Vision: Algorithms and Applications*. Springer, ISBN: 978-1-84882-935-0, 2010.
- [85] SZYDZIK, T., CALLICO, G. M., AND NUNEZ, A. Efficient fpga implementation of a high-quality super-resolution algorithm with real-time performance. *IEEE Transactions on Consumer Electronics* 57, 2 (May 2011), 664–672.
- [86] TIMOFTE, R., SMET, V. D., AND GOOL, L. V. Anchored neighborhood regression for fast example-based super-resolution. In *International Conference on Computer Vision (ICCV 2013)* (December 2013).
- [87] TOMASI, C., AND MANDUCHI, R. Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on* (1998), pp. 839–846.
- [88] USLUBAS, S., MAANI, E., AND KATSAGGELOS, A. K. A resolution adaptive video compression system. *Intelligent Multimedia Communication: techniques and applications, Springer* 280 (September 2010), 167–194.
- [89] VILLENA, S., VEGA, M., BABACAN, S., MOLINA, R., AND KATSAGGELOS, A. Bayesian combination of sparse and non-sparse priors in image super resolution. *Digital Signal Processing* 23, 2 (2013), 530 – 541.
- [90] VILLENA, S., VEGA, M., MOLINA, R., AND KATSAGGELOS, A. Bayesian super-resolution image reconstruction using an $l1$ prior. In *Image and Signal Processing and Analysis, 2009. ISPA 2009. Proceedings of 6th International Symposium on* (2009), pp. 152–157.
- [91] VOLKOV, V. Better Performance at Lower Occupancy, Sept. 2010.
- [92] WANG, R.-J., CHIENA, M.-C., AND CHANG, P.-C. Adaptive down-sampling video coding. In *Proceedings of SPIE, Multimedia on mobile devices, vol. 7542* (2010).
- [93] WANG, Z., BOVIK, A., SHEIKH, H., AND SIMONCELLI, E. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on* 13, 4 (2004), 600–612.

- [94] WERNER, M., STABERNACK, B., AND RIECHERT, C. Hardware implementation of a full hd real-time disparity estimation algorithm. *Consumer Electronics, IEEE Transactions on* 60, 1 (2014), 66–73.
- [95] XU, C., KIRK, S., AND JENKINS, S. Tiling for performance tuning on different models of gpus. In *Information Science and Engineering (ISISE), 2009 Second International Symposium on* (Dec 2009), pp. 500–504.
- [96] YANG, C., GOTO, S., AND IKENAGA, T. High performance vlsi architecture of fractional motion estimation in h.264 for hdtv. In *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on* (May 2006), pp. 4 pp.–.
- [97] YANG, D., SUN, J., LEE, J., LIANG, G., JENKINS, D. D., PETERSON, G. D., AND LI, H. Performance comparison of cholesky decomposition on gpus and fpgas. In *Symposium on Application Accelerators in High Performance Computing* (2010).
- [98] YANG, J., WRIGHT, J., HUANG, T., AND MA, Y. Image super-resolution via sparse representation. *Image Processing, IEEE Transactions on* 19, 11 (2010), 2861–2873.
- [99] ZENG, B., AND VENETSANOPOULOS, A. N. A JPEG-based interpolative image coding scheme. In *IEEE International Conference on Acoustics, Speech, and Signal Processing* (1993), pp. 393–396.
- [100] ZHANG, X., AND WU, X. Can lower resolution be better? In *IEEE Data Compression Conference* (2008), pp. 302–311.
- [101] ZHU, Y., ZHANG, Y., AND YUILLE, A. L. Single image super-resolution using deformable patches. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on* (June 2014), pp. 2917–2924.