



DISSERTAÇÃO

Mestrado em Engenharia Electrotécnica - Energia e Automação

Codificação Compatível de Vídeo 3D com o Algoritmo HEVC

GILBERTO JORGE

Leiria, Junho de 2014



DISSERTAÇÃO

Mestrado em Engenharia Electrotécnica - Energia e Automação

Codificação Compatível de Vídeo 3D com o Algoritmo HEVC

GILBERTO JORGE

Dissertação de Mestrado realizada sob a orientação do Doutor Sérgio M. M. de Faria, Professor da Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria e co-orientação do Doutor Nuno M. M. Rodrigues, Professor da Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria.

Leiria, Junho de 2014

À minha família.

Agradecimentos

Gostaria de agradecer ao meu orientador, Professor Sérgio Maciel de Faria, co-orientador, Professor Nuno Miguel Rodrigues, pela orientação, ensinamentos, apoio disponibilizado e amizade ao longo do desenvolvimento do presente trabalho.

Agradeço ao Instituto de Telecomunicações pela possibilidade de realização do trabalho de investigação que serviu de base a esta dissertação. Agradeço em especial à equipa da Delegação de Leiria por todas as condições proporcionadas, humanas e materiais, para a realização deste trabalho.

Aos meus colegas e amigos do Instituto de Telecomunicações e Escola Superior de Tecnologia e Gestão de Leiria, pelo apoio e amizade.

À minha família pelo seu apoio incondicional.

Este projeto foi financiado pela Fundação para a Ciência e Tecnologia (FCT), no âmbito do projeto PTDC/EEA-TEL/099387/2008 com o título “COMUVI-Compressão de vídeo multi-vista”.

Abstract

In this dissertation we study new and efficient methods to enhance the performance of compatible methods to transmit stereoscopic (3D) video sequences using a traditional 2D HEVC-based encoder and a simulcast transmission scenario.

Despite maintaining compatibility with the legacy 2D video equipments, the simulcast transmission requires a high bit rate. In the absence of adequate 3D coding tools, it is possible to reduce the bit rate by using asymmetric video compression, where the base view is encoded at full spatial resolution and the auxiliary view is transmitted with a lower spatial resolution and then up-sampled at the decoder.

We have developed a method for enhancing the up-sampled auxiliary view at the decoder, that uses detail information from the base view (with full resolution). This process relies on Affine transforms to perform a geometric mapping between the high frequency detail in the full resolution base view and the low resolution auxiliary one. Additionally, in order to attain image continuity among regions and avoid blocking artifacts, the algorithm maps the details from the base view using a triangular mesh as a reference.

The proposed technique is compared against the disparity block search method. The results show that, for some sequences, it improves not only the objective quality (PSNR) by up to 2.2 dB, but also the subjective quality.

Keywords: compatible 3D video coding, asymmetric stereoscopic video, inter-view enhancement

Resumo

Esta dissertação apresenta um trabalho sobre codificação de vídeo 3D compatível com vídeo 2D. Tem por base o desenvolvimento de um método para melhorar, no decodificador, a reconstrução de uma vista subamostrada resultante de uma transmissão *simulcast* usando a norma de codificação de vídeo H.265 (informalmente denominada de *High Efficiency Video Coding* (HEVC)).

Apesar de manter a compatibilidade com vídeo 2D a transmissão *simulcast* normalmente requer uma taxa de transmissão elevada. Na ausência de ferramentas de codificação 3D adequadas é possível reduzir a taxa de transmissão utilizando compressão assimétrica do vídeo, onde a vista base é codificada com a resolução espacial original, enquanto que a vista auxiliar é codificada com uma resolução espacial menor, sendo sobreamostrada no decodificador.

O método desenvolvido visa melhorar a vista auxiliar sobreamostrada no decodificador utilizando informação dos detalhes da vista base, ou seja, as componentes de alta frequência. Este processo depende de transformadas Afim para realizar um mapeamento geométrico entre a informação de alta frequência da vista base de resolução completa e a vista auxiliar de menor resolução. Adicionalmente, de modo a manter a continuidade do conteúdo da imagem entre regiões, evitando artefatos de blocos, o mapeamento utiliza uma malha de triangulação da vista auxiliar aplicado à imagem de detalhes obtida a partir da vista base.

A técnica proposta é comparada com um método de estimação de disparidade por correspondência de blocos, sendo que os resultados mostram que para algumas sequências a técnica desenvolvida melhora não só a qualidade objetiva (PSNR) até 2.2 dB, mas também a qualidade subjetiva, para a mesma taxa de compressão global.

Palavras-chave: codificação vídeo 3D compatível, vídeo estereoscópico assimétrico, melhoria entre vistas.

Conteúdo

Agradecimentos	iii
Abstract	v
Resumo	vii
Lista de Abreviaturas	xxi
1 Introdução	1
1.1 Contexto e motivação	1
1.2 Objetivos	2
1.3 Organização do documento	3
2 Codificação Assimétrica de Vídeo 3D com Métodos Compatíveis	5
3 HEVC	13
3.1 Introdução	13
3.2 Descrição Geral do Codificador	14
4 Codificação Compatível 3D	23
4.1 Codificação assimétrica	23
4.2 Método baseado em pesquisa de bloco	24
4.3 Método baseado em malha triangular	25
4.4 Adição da imagem de detalhes	30
5 Resultados Experimentais	35
5.1 Sequências de teste	36
5.2 Determinação de parâmetros	40
5.3 Imagem de referência na pesquisa de bloco	46
5.4 Adição da imagem de detalhes	48
5.5 Discussão de resultados	51
6 Conclusão e Trabalho Futuro	61

Bibliografia	65
A Testes de parâmetros	69
A.1 Tamanho do <i>kernel</i> para o método da malha triangular	69
A.2 Largura da janela de pesquisa para o método baseado na malha triangular	77
A.3 Tamanho do bloco para o método baseado na pesquisa de blocos	84
A.4 Largura da janela de pesquisa para o método baseado na pesquisa de blocos	92
A.5 Vista de referência na pesquisa de blocos	99
A.6 Avaliação dos métodos de adição de detalhes	113
A.7 Avaliação dos métodos de adição de detalhes - BD-PSNR	140
A.8 Avaliação dos métodos de melhoria da vista auxiliar	143
B Software de Simulação	155
B.1 Software de melhoria da vista auxiliar	155
C Publicações	159
C.1 Conferências	159

Lista de Figuras

2.1 Diagrama do codificador com subamostragem e predição compensada pela disparidade (adaptado de [7]).	7
2.2 Diagrama do método de remoção / reconstrução da crominância (adaptado de [9]).	9
2.3 Exemplo de padrão de filtragem (adaptado de [11]).	9
2.4 Imagens de referência entrecaladas nas restantes imagens subamostradas (adaptado de [13]).	10
2.5 Resultado obtido para o método da Super Resolução em vídeo 2D (adaptado de [13]).	11
2.6 Vídeo multivista com resolução assimétrica (adaptado de [14]).	11
3.1 Diagrama geral de blocos do codificador HEVC.	15
3.2 Exemplo simplificado de particionamento de uma imagem em CTU, CUs, PUs e TUs.	16
3.3 Direções angulares da predição Intra Angular de acordo com o tamanho da CU (adaptado de [18]).	17
3.4 Tipos de partição suportados na predição do HEVC (adaptado de [20]).	19

4.1	Diagrama geral de transmissão <i>simulcast</i> assimétrica.	23
4.2	Arquitetura geral do decodificador.	24
4.3	Pesquisa de bloco semelhante na vista auxiliar.	25
4.4	Diagrama do método da malha triangular.	26
4.5	Exemplo de triangulação de Delaunay (adaptado de [28]).	28
4.6	Triangulação de Delaunay aplicada à Imagem 1 da sequência <i>Balloons</i> . . .	28
4.7	Triângulo na vista auxiliar, aplicado na vista base e deslocado de acordo com a disparidade.	29
4.8	Método de adição seletiva de detalhes.	30
4.9	Exemplos do modo de decisão utilizando as imagens originais.	31
4.10	Exemplos com o modo de decisão utilizando a imagem de contornos. . . .	33
5.1	Sequências estéreo utilizadas nos testes, com resolução 1920×1088	37
5.2	Sequências estéreo utilizadas nos testes, com resolução 1280×960	37
5.3	Sequências estéreo utilizadas nos testes, com resolução 1024×768	38
5.4	Sequências estéreo utilizadas nos testes, com resolução 640×480	39
5.5	Exemplos de identificação da margem da vista auxiliar não existente na vista base.	41
5.6	Diferença de PSNR da vista auxiliar, das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar subamostrada na horizontal e vertical.	43
5.7	Diferença de PSNR da vista auxiliar, das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa no método da malha triangular, com a vista auxiliar subamostrada na horizontal e vertical.	45
5.8	Diferença de PSNR da sequência <i>Champagne Tower</i> em função da variação do tamanho do <i>kernel</i> , comparando a utilização da vista auxiliar e base como vista de referência na pesquisa do <i>kernel</i>	47
5.9	Comparação da taxa-distorção (RD) entre a adição direta e com base na máscara para a sequência <i>Poznan Hall 2</i> com a vista auxiliar subamostrada na horizontal.	49
5.10	Comparação da taxa-distorção (RD) entre a adição direta e com base na máscara para a sequência <i>GT Fly</i> com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.	50
5.11	Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência <i>GT Fly</i> com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.	52
5.12	Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência <i>Pantomime</i> com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.	53

5.13	Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência <i>Poznan Hall 2</i> com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.	54
5.14	Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência <i>Balloons</i> com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.	56
5.15	Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência <i>Kendo</i> com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.	57
5.16	Comparação subjetiva e objetiva num detalhe da Imagem 1 de <i>Book Arrival</i> (Cam. 10, Imagem 1) e diferenças linear e logarítmica entre as amostras obtidas e a amostra de referência.	58
5.17	Comparação subjetiva e objetiva num detalhe da Imagem 1 de <i>Book Arrival</i> (Cam. 10, Imagem 1) e diferenças linear e logarítmica entre as amostras obtidas e a amostra de referência.	60
A.1	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar subamostrada na horizontal (H).	69
A.2	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar subamostrada na vertical (V).	70
A.3	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar subamostrada na horizontal e vertical (H/V).	70
A.4	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H.	71
A.5	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar V.	71
A.6	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H/V.	72
A.7	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H.	72
A.8	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar V.	73
A.9	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H/V.	73
A.10	Diferença de PSNR da sequência <i>Book Arrival</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H.	74

A.11	Diferença de PSNR da sequência <i>Book Arrival</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar V.	74
A.12	Diferença de PSNR da sequência <i>Book Arrival</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H/V.	75
A.13	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H.	75
A.14	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar V.	76
A.15	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho do <i>kernel</i> , com a vista auxiliar H/V.	76
A.16	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	77
A.17	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	78
A.18	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	78
A.19	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	79
A.20	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	79
A.21	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	80
A.22	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	80
A.23	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	81
A.24	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	81
A.25	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	82
A.26	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	82
A.27	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	83

A.28	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do bloco, com a vista auxiliar H.	84
A.29	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do bloco, com a vista auxiliar V.	85
A.30	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho do bloco, com a vista auxiliar H/V.	85
A.31	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho do bloco, com a vista auxiliar H.	86
A.32	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho do bloco, com a vista auxiliar V.	86
A.33	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho do bloco, com a vista auxiliar H/V.	87
A.34	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho do bloco, com a vista auxiliar H.	87
A.35	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho do bloco, com a vista auxiliar V.	88
A.36	Diferença de PSNR das sequências <i>Balloons</i> e <i>Kendo</i> em função da variação do tamanho do bloco, com a vista auxiliar H/V.	88
A.37	Diferença de PSNR da sequência <i>Book Arrival</i> em função da variação do tamanho do bloco, com a vista auxiliar H.	89
A.38	Diferença de PSNR da sequência <i>Book Arrival</i> em função da variação do tamanho do bloco, com a vista auxiliar V.	89
A.39	Diferença de PSNR da sequência <i>Book Arrival</i> em função da variação do tamanho do bloco, com a vista auxiliar H/V.	90
A.40	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho do bloco, com a vista auxiliar H.	90
A.41	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho do bloco, com a vista auxiliar V.	91
A.42	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho do bloco, com a vista auxiliar H/V.	91
A.43	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	92
A.44	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	93
A.45	Diferença de PSNR das sequências <i>GT Fly</i> e <i>Poznan Hall 2</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	93

A.46	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	94
A.47	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	94
A.48	Diferença de PSNR das sequências <i>Pantomime</i> e <i>Champagne Tower</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	95
A.49	Diferença de PSNR das sequências <i>Balloons</i> , <i>Kendo</i> e <i>Book Arrival</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	95
A.50	Diferença de PSNR das sequências <i>Balloons</i> , <i>Kendo</i> e <i>Book Arrival</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	96
A.51	Diferença de PSNR das sequências <i>Balloons</i> , <i>Kendo</i> e <i>Book Arrival</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	96
A.52	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.	97
A.53	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.	97
A.54	Diferença de PSNR das sequências <i>Ballroom</i> e <i>Exit</i> em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.	98
A.55	Seleção da vista de referência, no método da malha, para a sequência <i>GT Fly</i> , com a vista auxiliar H.	99
A.56	Seleção da vista de referência, no método da malha, para a sequência <i>GT Fly</i> , com a vista auxiliar V.	100
A.57	Seleção da vista de referência, no método da malha, para a sequência <i>GT Fly</i> , com a vista auxiliar H/V.	100
A.58	Seleção da vista de referência, no método da malha, para a sequência <i>Poznan Hall 2</i> , com a vista auxiliar H.	101
A.59	Seleção da vista de referência, no método da malha, para a sequência <i>Poznan Hall 2</i> , com a vista auxiliar V.	101
A.60	Seleção da vista de referência, no método da malha, para a sequência <i>Poznan Hall 2</i> , com a vista auxiliar H/V.	102
A.61	Seleção da vista de referência, no método da malha, para a sequência <i>Pantomime</i> , com a vista auxiliar H.	102

A.62	Seleção da vista de referência, no método da malha, para a sequência <i>Pantomime</i> , com a vista auxiliar V.	103
A.63	Seleção da vista de referência, no método da malha, para a sequência <i>Pantomime</i> , com a vista auxiliar H/V.	103
A.64	Seleção da vista de referência, no método da malha, para a sequência <i>Champagne Tower</i> , com a vista auxiliar H.	104
A.65	Seleção da vista de referência, no método da malha, para a sequência <i>Champagne Tower</i> , com a vista auxiliar V.	104
A.66	Seleção da vista de referência, no método da malha, para a sequência <i>Balloons</i> , com a vista auxiliar H.	105
A.67	Seleção da vista de referência, no método da malha, para a sequência <i>Balloons</i> , com a vista auxiliar V.	105
A.68	Seleção da vista de referência, no método da malha, para a sequência <i>Balloons</i> , com a vista auxiliar H/V.	106
A.69	Seleção da vista de referência, no método da malha, para a sequência <i>Kendo</i> , com a vista auxiliar H.	106
A.70	Seleção da vista de referência, no método da malha, para a sequência <i>Kendo</i> , com a vista auxiliar V.	107
A.71	Seleção da vista de referência, no método da malha, para a sequência <i>Kendo</i> , com a vista auxiliar H/V.	107
A.72	Seleção da vista de referência, no método da malha, para a sequência <i>Book Arrival</i> , com a vista auxiliar H.	108
A.73	Seleção da vista de referência, no método da malha, para a sequência <i>Book Arrival</i> , com a vista auxiliar V.	108
A.74	Seleção da vista de referência, no método da malha, para a sequência <i>Book Arrival</i> , com a vista auxiliar H/V.	109
A.75	Seleção da vista de referência, no método da malha, para a sequência <i>Ballroom</i> , com a vista auxiliar H.	109
A.76	Seleção da vista de referência, no método da malha, para a sequência <i>Ballroom</i> , com a vista auxiliar V.	110
A.77	Seleção da vista de referência, no método da malha, para a sequência <i>Ballroom</i> , com a vista auxiliar H/V.	110
A.78	Seleção da vista de referência, no método da malha, para a sequência <i>Exit</i> , com a vista auxiliar H.	111
A.79	Seleção da vista de referência, no método da malha, para a sequência <i>Exit</i> , com a vista auxiliar V.	111
A.80	Seleção da vista de referência, no método da malha, para a sequência <i>Exit</i> , com a vista auxiliar H/V.	112

A.81	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>GT Fly</i>	113
A.82	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>GT Fly</i>	114
A.83	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Poznan Hall 2</i>	114
A.84	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Poznan Hall 2</i>	115
A.85	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Pantomime</i>	115
A.86	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Pantomime</i>	116
A.87	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Pantomime</i>	116
A.88	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Champagne Tower</i>	117
A.89	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Champagne Tower</i>	117
A.90	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Champagne Tower</i>	118
A.91	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Balloons</i>	118
A.92	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Balloons</i>	119
A.93	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Balloons</i>	119
A.94	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Kendo</i>	120
A.95	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Kendo</i>	120
A.96	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Kendo</i>	121
A.97	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Book Arrival</i>	121
A.98	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Book Arrival</i>	122
A.99	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Book Arrival</i>	122

A.100	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Ballroom</i>	123
A.101	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Ballroom</i>	123
A.102	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Ballroom</i>	124
A.103	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência <i>Exit</i>	124
A.104	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência <i>Exit</i>	125
A.105	Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência <i>Exit</i>	125
A.106	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>GT Fly</i>	126
A.107	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>GT Fly</i>	127
A.108	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>GT Fly</i> . . .	127
A.109	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Poznan Hall 2</i> . 128	
A.110	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Poznan Hall 2</i> . 128	
A.111	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Poznan Hall 2</i>	129
A.112	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Pantomime</i> . .	129
A.113	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Pantomime</i> . .	130
A.114	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Pantomime</i> . 130	
A.115	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Champagne Tower</i>	131
A.116	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Champagne Tower</i>	131

A.117	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Champagne Tower</i>	132
A.118	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Balloons</i>	132
A.119	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Balloons</i>	133
A.120	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Balloons</i> . .	133
A.121	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Kendo</i>	134
A.122	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Kendo</i>	134
A.123	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Kendo</i> . . .	135
A.124	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Book Arrival</i> . .	135
A.125	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Book Arrival</i> . .	136
A.126	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Book Arrival</i> . .	136
A.127	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Ballroom</i> . . .	137
A.128	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Ballroom</i> . . .	137
A.129	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Ballroom</i> . .	138
A.130	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência <i>Exit</i>	138
A.131	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência <i>Exit</i>	139
A.132	Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência <i>Exit</i>	139
A.133	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>GT Fly</i> , com vista auxiliar H.	143
A.134	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>GT Fly</i> , com vista auxiliar V.	144

A.135	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Poznan Hall 2</i> , com vista auxiliar H.	144
A.136	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Poznan Hall 2</i> , com vista auxiliar V.	145
A.137	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Pantomime</i> , com vista auxiliar H.	145
A.138	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Pantomime</i> , com vista auxiliar V.	146
A.139	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Champagne Tower</i> , com vista auxiliar H.	146
A.140	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Champagne Tower</i> , com vista auxiliar V.	147
A.141	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Champagne Tower</i> , com vista auxiliar H/V.	147
A.142	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Balloons</i> , com vista auxiliar H.	148
A.143	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Balloons</i> , com vista auxiliar V.	148
A.144	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Kendo</i> , com vista auxiliar H.	149
A.145	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Kendo</i> , com vista auxiliar V.	149
A.146	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Book Arrival</i> , com vista auxiliar H.	150
A.147	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Book Arrival</i> , com vista auxiliar V.	150
A.148	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Book Arrival</i> , com vista auxiliar H/V.	151
A.149	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Ballroom</i> , com vista auxiliar H.	151
A.150	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Ballroom</i> , com vista auxiliar V.	152
A.151	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Ballroom</i> , com vista auxiliar H/V.	152
A.152	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Exit</i> , com vista auxiliar H.	153
A.153	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Exit</i> , com vista auxiliar V.	153

A.154	Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência <i>Exit</i> , com vista auxiliar H/V.	154
-------	---	-----

Lista de Tabelas

5.1	Sequências estéreo testadas.	36
5.2	Intervalo de teste do tamanho do <i>kernel</i> e bloco (<i>pixels</i>).	42
5.3	Intervalo de teste da largura da janela de pesquisa (<i>pixels</i>).	44
5.4	Resultados obtidos para o tamanho de bloco e janela de pesquisa (<i>pixels</i>).	45
5.5	Avaliação do método de adição de detalhes do método da malha triangular, com vista auxiliar subamostrada na horizontal e vertical.	50
5.6	Avaliação dos métodos de melhoria da vista auxiliar subamostrada na horizontal (H).	54
5.7	Avaliação dos métodos de melhoria da vista auxiliar subamostrada na vertical (V).	55
5.8	Avaliação dos métodos de melhoria da vista auxiliar subamostrada na horizontal e vertical (H/V).	55
A.1	Avaliação das técnicas de adição de detalhes no método da malha triangular, com vista auxiliar H.	140
A.2	Avaliação das técnicas de adição de detalhes no método da malha triangular, com vista auxiliar V.	140
A.3	Avaliação das técnicas de adição de detalhes no método da pesquisa de blocos, com vista auxiliar H.	141
A.4	Avaliação das técnicas de adição de detalhes no método da pesquisa de blocos, com vista auxiliar V.	141
A.5	Avaliação das técnicas de adição de detalhes no método da pesquisa de blocos, com vista auxiliar H/V.	142

Lista de Abreviaturas

2D	Bidimensional
3D	Tridimensional
AMP	<i>Asymmetric Motion Partition</i> (Partição Assimétrica de Movimento)
AVC	<i>Advanced Video Coding</i> (Codificação de Vídeo Avançada)
BD-PSNR	<i>Bjontegaard Delta PSNR</i>
BR	<i>Bjontegaard Delta Bit Rate</i> (<i>Delta Bjontegaard</i> de Taxa de Transmissão)
CABAC	<i>Context Adaptive Binary Arithmetic Coding</i> (Codificação Aritmética Binária com Adaptação ao Contexto)
CB	<i>Coding Block</i> (Bloco de Codificação)
CTB	<i>Coding Tree Block</i> (Bloco da Árvore de Codificação)
CTU	<i>Coding Tree Unit</i> (Unidade da Árvore de Codificação)
CU	<i>Coding Unit</i> (Unidade de Codificação)
CfP	<i>Call for Proposals</i> (Apresentação de Propostas)
Codec	Codificador / Decodificador
DBF	<i>Deblocking Filter</i> (Filtro de <i>Deblocking</i>)
DCP	<i>Disparity Compensated Prediction</i> (Predição por Compensação de Disparidade)
DCT	<i>Discrete Cosine Transform</i> (Transformada Discreta de Cosseno)
DST	<i>Discrete Sine Transform</i> (Transformada Discreta de Seno)
FCT	Fundação para a Ciência e Tecnologia

GCC	<i>GNU Compiler Collection</i> (Coleção de Compiladores GNU)
GOP	<i>Group of Pictures</i> (Grupo de Imagens)
HEVC	<i>High Efficiency Video Coding</i> (Codificação de Vídeo de Elevada Eficiência)
HM	<i>HEVC Test Model</i> (Modelo de Teste do HEVC)
HP	<i>High Profile</i> (Perfil <i>High</i>)
HVS	<i>Human Visual System</i> (Sistema Visual Humano)
IEC	<i>International Electrotechnical Commission</i> (Comissão Eletrotécnica Internacional)
ISO	<i>International Standards Organization</i> (Organização Internacional de Normas)
ITU-T	<i>International Telecommunication Union - Telecommunication Standardization Sector</i> (União Internacional de Telecomunicações - Setor de Normas em Telecomunicações)
JCT-VC	<i>Joint Collaborative Team on Video Coding</i> (Equipa de Colaboração Conjunta em Codificação de Vídeo)
JSVM	<i>Joint Scalable Video Model</i> (Modelo de Vídeo Escalável Conjunto)
MCP	<i>Motion Compensated Prediction</i> (Predição por Compensação de Movimento)
MC	<i>Motion Compensation</i> (Compensação de Movimento)
MPEG	<i>Moving Picture Experts Group</i> (Grupo de Especialistas de Imagens em Movimento)
MP	<i>Main Profile</i> (Perfil <i>Main</i>)
MSE	<i>Mean Square Error</i> (Erro Quadrático Médio)
MV	<i>Motion Vector</i> (Vetor de Movimento)
OpenCV	<i>Open Source Computer Vision Library</i> (Biblioteca de Funções de Visão Computacional em Código Aberto)
PB	<i>Prediction Block</i> (Bloco de Predição)

PSNR	<i>Peak Signal-to-Noise Ratio</i> (Pico de Relação de Sinal-Ruído)
PU	<i>Prediction Unit</i> (Unidade de Predição)
QP	<i>Quantization Parameter</i> (Parâmetro de Quantificação)
RD	<i>Rate-Distortion</i> (Taxa-Distorção)
SAD	<i>Sum of Absolute Differences</i> (Soma Absoluta das Diferenças)
SAO	<i>Sample Adaptive Offset</i> (<i>Offset</i> Adaptativo de Amostra)
SVC	<i>Scalable Video Coding</i> (Codificação de Vídeo Escalável)
TB	<i>Transform Block</i> (Bloco de Transformada)
TMuC	<i>Test Model under Consideration</i> (Modelo de Teste em Consideração)
TU	<i>Transform Unit</i> (Unidade de Transformada)
VCEG	<i>Video Coding Experts Group</i> (Grupo de Especialistas em Codificação de Vídeo)
WPP	<i>Wavefront Parallel Processing</i> (Processamento em Paralelo <i>Wavefront</i>)

Capítulo 1

Introdução

1.1 Contexto e motivação

A utilização de sequências de vídeo e imagens tridimensionais (3D) tem suscitado recentemente um grande entusiasmo por parte da comunidade científica e industrial. Este entusiasmo é justificado pela melhoria do efeito envolvente que a ilusão de profundidade provoca no utilizador, refletindo-se em produtos e serviços de multimédia que merecem cada vez mais a atenção do consumidor final.

A utilização de várias vistas da mesma cena para obtenção de vídeo 3D apresenta no entanto alguns desafios importantes, como o aumento significativo de necessidade de largura de banda, quando comparado com a utilização de uma única vista ou a compatibilidade dos sistemas 3D com os sistemas tradicionais, fortemente implantados no mercado.

O método tradicional utilizado para reduzir os requisitos de largura de banda baseia-se em técnicas de redução de redundância entre vistas, assim como exploração da redundância espacial e temporal da sequência [1].

Contudo, podem ainda ser exploradas outras características do Sistema Visual Humano (HVS – *Human Visual System*) úteis na visualização de vídeo estereoscópico, utilizando, por exemplo, a codificação assimétrica de vídeo, i.e., a compressão das vistas da sequência com diferentes níveis de qualidade.

Por exemplo, se uma das vistas fornecer um nível de qualidade suficientemente elevado (vista base), a outra vista (auxiliar) pode ser codificada com qualidade inferior [2], uma vez que o cérebro irá fundir essas imagens possibilitando a percepção de profundidade [3]. No entanto, existem casos em que essa degradação é assinalável, ou tem um impacto negativo devido à dificuldade de percepção de pequenas diferenças de profundidade. Nesse caso uma possível solução poderá passar pela obtenção dos detalhes perdidos pela vista

de qualidade inferior, através da informação existente na vista base.

Uma técnica que permite a redução da taxa de transmissão consiste no envio em simultâneo (*simulcast*) de vídeo estéreo, baseado na codificação assimétrica, codificando uma vista com a resolução original e a outra com metade da resolução (na direção horizontal e/ou vertical). Após a descodificação e sobreamostragem da vista auxiliar, esta apresenta uma qualidade objetiva consideravelmente inferior, relativamente à vista base, devido à perda de dados de forma permanente durante o processo de subamostragem antes da codificação.

Porém, atendendo às semelhanças entre vistas, parte dos detalhes em falta podem estar presentes na vista base, que poderá ser utilizada para melhorar a vista auxiliar. Mas, este processo não pode ser efetuado de forma direta, devido à existência de um deslocamento espacial (disparidade) entre as duas vistas, que depende da profundidade dos objetos da cena em relação às câmaras e da distância/ângulo entre os eixos óticos das câmaras. A abordagem convencional a este processo utiliza métodos baseados em pesquisa de blocos para compensar esta disparidade. No entanto, estes métodos apresentam desvantagens consideráveis relativamente ao efeito de bloco produzido e dependência da estimativa da disparidade, que por vezes não produz o resultado esperado.

O método proposto pretende evitar o efeito de bloco e melhorar a continuidade entre regiões, aplicando uma partição triangular à imagem auxiliar, adaptando-a à imagem de resolução original, através de um mapeamento de disparidade entre os vértices dos triângulos de ambas as vistas. A adaptação é efetuada com transformadas Afim, sendo os detalhes do conteúdo da imagem da vista base processados e adicionados à imagem de baixa resolução. É efetuada a comparação entre este método e um método baseado em pesquisa de blocos (*block-matching*), que divide a imagem em blocos de tamanho fixo e usa a estimação de disparidade do bloco para adicionar os detalhes correspondentes.

Ambas as vistas, de resolução original e baixa resolução, foram codificadas com o software de referência (HM 13.0) [4] da norma H.265/*High Efficiency Video Coding* (HEVC) [5].

1.2 Objetivos

Inicialmente o trabalho envolveu o estudo do projeto da norma do HEVC e a realização de testes de codificação e descodificação. Seguidamente foram desenvolvidos os seguintes tópicos:

- Realização da codificação assimétrica de sequências de vídeo estéreo, onde a vista base é codificada com resolução original e a vista auxiliar é subamostrada na direção horizontal e/ou vertical para metade do respetivo valor original, antes da codificação;

- Tendo em conta a semelhança entre as vistas de sequências estéreo, pretende-se melhorar a vista auxiliar após a descodificação, utilizando informação de alta frequência disponível na vista base. Neste sentido desenvolveu-se um método para adaptação dos detalhes (alta frequência) do conteúdo da vista base à vista auxiliar, que não dependa exclusivamente da estimação de disparidade por pesquisa de blocos entre as imagens de cada vista;
- Efetuar a adição dos detalhes (alta frequência) da vista base à vista auxiliar com a informação da adaptação de conteúdo entre as vistas;
- Comparação dos resultados obtidos do método desenvolvido com os resultados obtidos usando um método de estimação de disparidade por pesquisa de blocos;
- Analisar o desempenho obtido usando o método proposto.

1.3 Organização do documento

Após a descrição geral do trabalho e dos objetivos, o Capítulo 2 apresenta o estado da arte na codificação compatível 3D utilizando vídeo estereoscópico assimétrico. No Capítulo 3 é realizado um estudo do codificador / decodificador (codec) HEVC. O Capítulo 4 mostra os algoritmos utilizados para a melhoria das sequências subamostradas na codificação assimétrica, de modo a obter a codificação de vídeo 3D num modo compatível com 2D, minimizando a perda de qualidade numa das vistas. Os resultados são disponibilizados no Capítulo 5, onde é efetuada a descrição dos testes realizados e análise/comparação dos resultados obtidos. O Capítulo 6 apresenta as conclusões e trabalho futuro.

No Apêndice A podem ser consultados os resultados obtidos sob a forma de gráficos de desempenho taxa-distorção (RD – *Rate-Distortion*) e tabelas com os valores calculados com a métrica *Delta Bjøntegaard* PSNR (BD-PSNR) e percentagem de *Delta Bjøntegaard* de taxa de transmissão (%BR – *Bjøntegaard Delta Bit Rate*). A descrição e o guia de utilização do software desenvolvido encontra-se no Apêndice B. No Apêndice C é indicado o artigo de conferência realizado com o desenvolvimento do presente trabalho.

Capítulo 2

Codificação Assimétrica de Vídeo 3D com Métodos Compatíveis

A utilização de tecnologia com vídeo num formato tridimensional (3D) ganha cada vez mais popularidade devido ao efeito envolvente que a percepção de profundidade provoca no espectador. Algumas ferramentas de compressão de vídeo tentam tirar partido das semelhanças entre as várias vistas de vídeo 3D através da redundância entre vistas, para além de técnicas que se encontram tradicionalmente na codificação de vídeo bidimensional (2D), nomeadamente, a exploração da redundância espacial e temporal da sequência. No entanto, a capacidade de percepção estereoscópica do ser humano permite que uma das vistas de uma sequência estéreo possa apresentar menor qualidade que a outra, criando assim a possibilidade de desenvolvimento de métodos que explorem esta característica.

O estado da arte relativo ao desenvolvimento de métodos ligados ao vídeo estereoscópico assimétrico, mostra que se trata de uma técnica com elevado potencial de utilização prática, sendo cada vez mais estudada. No entanto, não se observa a mesma atividade no que diz respeito ao desenvolvimento de métodos que contribuam para a melhoria da vista de baixa resolução com o intuito de melhorar a percepção de profundidade. Esta observação pode ser compreendida no âmbito da vasta quantidade de potenciais beneficiários deste tipo de tecnologia, nomeadamente na área do entretenimento, onde apenas se procura um grau de percepção de profundidade suficiente para uma experiência imersiva. A possibilidade de obtenção de uma percepção de profundidade com maior exatidão, tem em princípio um conjunto de destinatários em áreas específicas, onde o controlo fino de profundidade é crítico. Por exemplo, sistemas de visão artificial ou operadores laboratoriais / industriais com necessidade de controlar equipamento através de vídeo estereoscópico.

Perkins [3] foi um dos primeiros autores a introduzir o conceito de utilização de resoluções diferentes em pares de imagens estéreo. Ao desfocar o conteúdo de uma das vistas através da utilização de um filtro passa-baixo, obteve uma imagem de baixa resolução e

outra de elevada resolução. Os resultados mostraram que era possível reduzir a taxa de transmissão em 46% (ou mais, no caso de codificação da imagem de resolução original), relativamente a um sistema sem compressão, sendo observada uma reduzida degradação subjetiva na qualidade da imagem.

Outra técnica utilizada foi a codificação preditiva no domínio da transformada compensada pela disparidade, que tenta minimizar o erro quadrático médio (MSE – *Mean Square Error*) entre o par estéreo original e o codificado [3]. Neste caso, após a codificação independente de uma das vistas e divisão da outra imagem em blocos quadrados, para cada bloco é efetuada a aproximação ao bloco da imagem codificada, sendo a respetiva identificação transmitida ao recetor. A transformada do bloco é predita com base na transformada do bloco aproximado, sendo no final o erro de predição no domínio da transformada, quantificado e transmitido ao recetor. Neste caso foi obtida uma redução de 21%, relativamente à codificação independente da imagem.

Stelmach *et al.* [2] aplicou níveis de compressão diferentes a cada vista usando essencialmente três métodos: a quantificação baseada em transformada discreta de cosseno (DCT – *Discrete Cosine Transform*); filtragem passa-baixo; e uma combinação de filtragem e quantificação. Verificou-se que, com a filtragem passa-baixo, a perceção binocular era dominada pela imagem de elevada qualidade, enquanto que para a quantificação correspondia aproximadamente à média das imagens em cada olho. Os resultados mostraram que o sistema visual humano (HVS – *Human Visual System*) efetua a ponderação dos dados visualizados em cada olho dependendo do tipo de degradação, verificando-se que alterações na resolução espacial produzem imagens de melhor qualidade que alterações na quantificação. Com a combinação de ambos os métodos, a perceção de profundidade é dominada pela quantificação.

A subamostragem temporal e/ou espacial de uma das vistas foi analisada em [6]. Para efetuar o redimensionamento espacial a vista auxiliar é subamostrada antes da codificação. A subamostragem consiste na decimação e aplicação de um filtro passa-baixo para limitar o *aliasing*. Este processo é aplicado à luminância e crominâncias na direção horizontal e vertical. Para efetuar a subamostragem temporal a vista auxiliar é decimada no tempo, removendo determinadas imagens em intervalos regulares ao longo da sequência. Com o método desenvolvido concluiu-se que vídeos estereoscópicos podem ser comprimidos a uma taxa de cerca de 1.2 vezes em relação a de vídeos monoscópicos, com reduzida degradação da qualidade.

Em [7] a subamostragem espacial é aplicada à imagem da vista auxiliar (neste caso corresponde à vista direita), antes da predição entre vistas, combinando os benefícios da predição compensada pela disparidade e codificação assimétrica. A estrutura de predição do codificador / decodificador (codec) desenvolvido, consiste na compressão da vista

vista auxiliar utiliza um quarto da resolução da vista base (metade da resolução em cada uma das direções da imagem). Tal como na codificação de vídeo simétrico, é efetuada a predição da vista base para a vista auxiliar, de modo a diminuir ainda mais a largura de banda desta vista. Utiliza-se um processo em que a predição é efetuada diretamente sobre a imagem base considerando apenas parte das amostras da imagem, tal como sucederia se a subamostragem fosse realizada. Assim minimiza-se a utilização de memória adicional e eventualmente a complexidade de processamento.

As condições inerentes à obtenção de conteúdo através de vídeo multivista, tais como calibração e parâmetros não ideais das câmaras, podem-se traduzir numa menor eficiência da predição entre vistas [8]. A eficiência é ainda mais reduzida devido à utilização de apenas um quarto do número inteiro de pixels na vista base para efetuar a predição da vista auxiliar. A existência de regiões com diferentes níveis de profundidade poderá resultar numa imagem desfocada em algumas zonas devido ao desajuste de focalização, contribuindo também para uma menor eficiência da predição entre vistas. Justifica-se assim o desenvolvimento de um filtro adaptativo por regiões da imagem, já que um filtro global não seria tão eficiente. Os filtros são gerados assumindo a existência de um determinado número de níveis de profundidade, sendo proposto um método para segmentar a imagem em diferentes regiões de profundidade através do algoritmo K-means, agrupando os vetores de disparidade em classes (obtidos através da localização de bloco correspondente, *block matching*). Neste caso, observou-se uma poupança até 27% de taxa de transmissão comparativamente aos métodos sem a filtragem adaptativa [8].

Em [9] a vista auxiliar de menor resolução foi obtida através da degradação da componente de luminância de uma das vistas, remoção da informação das componentes de crominância e explorando as correlações entre vistas. A vista base é codificada de forma independente, podendo ser descodificada do mesmo modo, não necessitando de um processamento específico. As componentes de crominância da vista auxiliar são depois reconstruídas no descodificador a partir de informação existente nas componentes de luminância e crominância da vista base. Este processo utiliza a estimação de disparidade baseada em blocos na componente de luminância entre as duas vistas no descodificador. O diagrama do método descrito é ilustrado na Figura 2.2. Desta forma foi possível obter uma qualidade subjetiva similar à qualidade da sequência de vídeo original, mas com uma significativa redução de taxa de transmissão (15% a 35%).

Caso o vídeo estéreo assimétrico seja visualizado por pessoas com um olho dominante, não correspondente à vista de maior resolução, a qualidade global percebida não irá ser semelhante à sequência com resolução original [10]. Como os testes de entrecalamento temporal de vista de baixa resolução com alta resolução, resultaram numa comutação perceptível não desejável, em [11] foi sugerido a aplicação de um filtro passa-baixo (filtro Gaussiano) a *slices* horizontais em ambas as vistas, enquanto que a *slice* correspondente

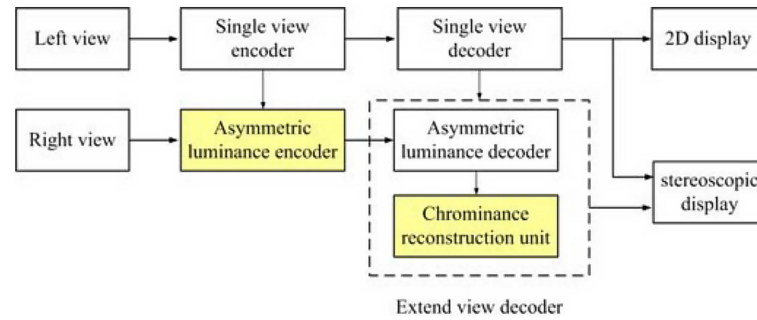


Figura 2.2: Diagrama do método de remoção / reconstrução da crominância (adaptado de [9]).

na outra vista não seria alterada (Figura 2.3). Os resultados obtidos mostraram que a utilização deste método, com a suavização nos contornos das *slices*, é uma técnica eficiente para vídeo estereoscópico assimétrico, permitindo manter a qualidade de percepção semelhante ao vídeo original, com uma redução significativa de taxa de transmissão.

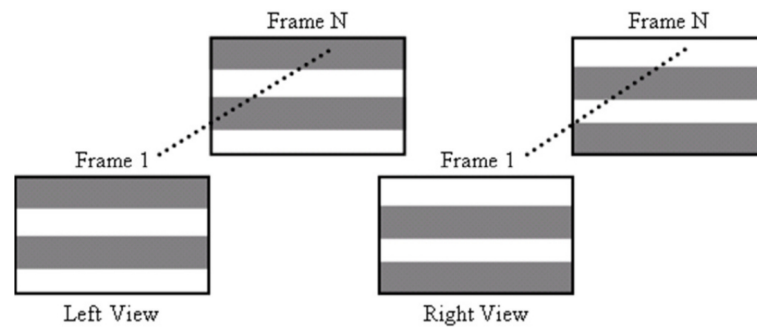


Figura 2.3: Exemplo de padrão de filtragem (adaptado de [11]).

Em [12] é proposto um esquema de melhoria de imagem de baixa resolução através de dois métodos de regularização. Primeiro, através de uma função de regularização baseada em saliências, adaptativa espacialmente, proposta para ajustar o grau de regularização em diferentes regiões da imagem. Depois, é proposto um esquema de regularização sensível a oclusões de modo a limitar problemas derivados da estimação errada de disparidade. Assim, o esquema de regularização sensível a oclusões explora a correlação entre a imagem de baixa resolução e a imagem de alta resolução observada na vista vizinha e, adaptativamente, marca as regiões onde existe oclusão de modo a que não contribuam para a estimação da imagem de resolução original desconhecida. Os resultados experimentais permitiram concluir que os métodos propostos em conjunto são superiores às abordagens convencionais usadas nos testes comparativos efetuados.

Em [13] é utilizado um método para vídeo 2D que implica que as imagens de referência de resolução original, em menor número, sejam intercaladas nas restantes imagens subamostradas. As imagens de resolução original são usadas para ajudar a recuperação de informação de alta frequência das restantes imagens sobreamostradas na desco-

dificação. Este processo é possível devido à elevada correlação existente entre imagens temporalmente próximas. A Figura 2.4 ilustra as imagens de referência (*key frames*) e subamostradas (*non-key frames*) num intervalo de uma sequência de vídeo.

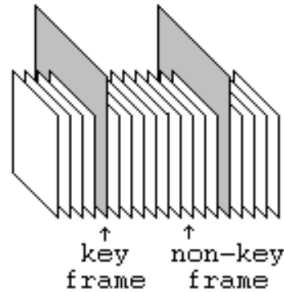


Figura 2.4: Imagens de referência entrecaladas nas restantes imagens subamostradas (adaptado de [13]).

O método proposto para a recuperação da informação de alta frequência das imagens sobreamostradas utiliza a estimação de movimento para encontrar a melhor correspondência entre um bloco de uma imagem sobreamostrada e uma de referência. A estimação de movimento das imagens subamostradas baseia-se nas imagens de referência. Na decodificação, a estimação de movimento em vez de utilizar a versão subamostrada de uma imagem de referência, é aplicado um filtro passa-alto a esta versão subamostrada, bem como à imagem que se pretende recuperar, realizando depois a pesquisa de correspondências. De acordo com os autores, esta informação de passa-banda permite efetuar uma melhor predição das bandas de alta frequência.

Com a obtenção da melhor correspondência, a informação de alta frequência do bloco da imagem de referência é adicionado ao bloco da imagem sobreamostrada. No entanto, a melhor correspondência geralmente não é perfeita, gerando ruído na imagem que se pretende melhorar. Deste modo, em vez de adicionar toda a informação, a alta frequência adicionada é o resultado da convolução da imagem de referência original com um *kernel* passa-alto. Os resultados obtidos evidenciaram melhorias obtidas nas imagens recuperadas relativamente à utilização de apenas interpolação. A Figura 2.5 ilustra, para uma sequência, os resultados obtidos da qualidade objetiva em função da frequência de imagens de referência. Aqui observa-se a comparação dos resultados utilizando apenas técnicas clássicas de sobreamostragem (bilinear e bicúbica) e a sobreamostragem com o algoritmo desenvolvido para vários tamanhos de bloco.

Em [14], um método de Super Resolução é utilizado em vídeo multivista. Neste tipo de vídeo são utilizadas imagens de textura e mapas de profundidade para gerar vistas adicionais, não se tratando portanto de um método que possa ser aplicado ao vídeo estereoscópio assimétrico de duas vistas de textura. Este método tem o objetivo de melhorar as vistas de baixa resolução usando o conteúdo de alta frequência proveniente

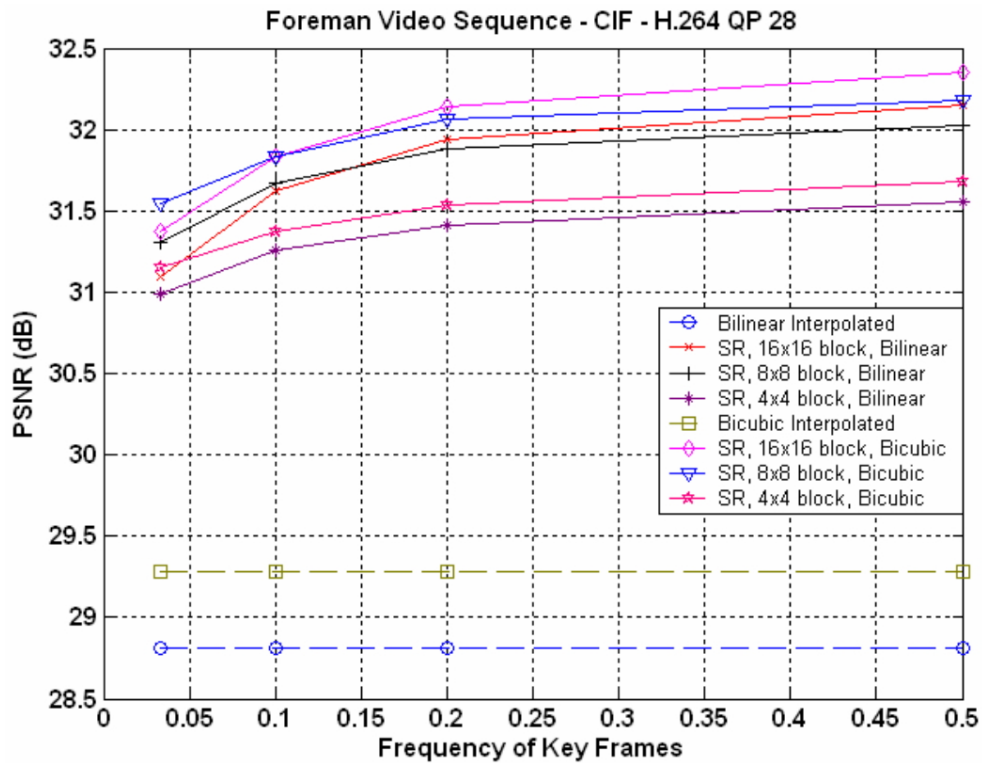


Figura 2.5: Resultado obtido para o método da Super Resolução em vídeo 2D (adaptado de [13]).

das vistas adjacentes de resolução original. A Figura 2.6 ilustra um exemplo de vídeo de múltiplas vistas em resolução mista com mapas de profundidade. Este método tem como finalidade colmatar as desvantagens significativas ao adaptar vídeo de resolução mista a outras funções baseadas na utilização de múltiplas vistas, tal como a síntese de vistas intermédias.

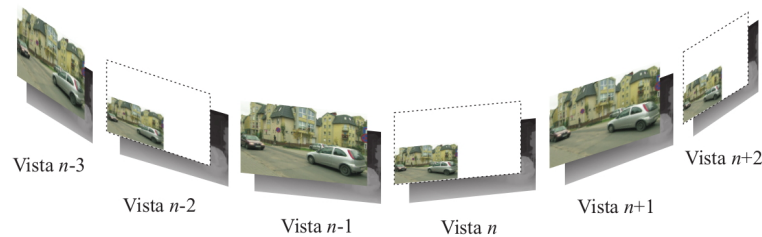


Figura 2.6: Vídeo multivista com resolução assimétrica (adaptado de [14]).

Tal como em [13], as vistas de referência com resolução original são utilizadas para restaurar regiões das restantes vistas auxiliares subamostradas. O método proposto assume que qualquer vista de resolução original pode ser decomposta em componentes de baixa e alta frequência. Para obter a estimativa da versão original de uma das vistas de baixa resolução, deve ser encontrada a estimativa de alta frequência que somada ao conteúdo de baixa resolução forma a imagem recuperada. A estimativa da alta frequência é obtida em

várias etapas. Inicialmente é realizada a projeção da vista adjacente de resolução original na direção da n -ésima vista com auxílio dos mapas de profundidade. O resultado é subamostrado e sobreamostrado para obter a versão passa-baixo. A imagem com a informação de alta frequência é obtida através da subtração da imagem da primeira etapa com a da segunda etapa. De acordo com os resultados obtidos, verificam-se ganhos resultantes do método desenvolvido em termos de qualidade objetiva e subjetiva em relação às imagens sintetizadas.

Capítulo 3

HEVC

3.1 Introdução

Com o intuito de responder à persistente procura de utilização de conteúdos de vídeo, bidimensional (2D) e tridimensional (3D), com cada vez mais resolução (espacial e temporal), surge a necessidade de melhorar a eficiência de compressão de vídeo. Durante o desenvolvimento deste trabalho foi utilizado a norma H.265/*High Efficiency Video Coding* (HEVC) no processo de codificação e descodificação de vídeo estereoscópico. O método desenvolvido atua apenas no descodificador como ferramenta opcional para melhoria de vídeo estéreo assimétrico, tratando-se assim de um método compatível com a norma utilizada. Este capítulo descreve, de uma forma geral, o funcionamento do HEVC de modo a compreender o processo de compressão aplicado à informação de vídeo utilizado com o método proposto.

No seguimento do estudo de técnicas que permitissem a melhoria da norma de compressão de vídeo H.264/MPEG-4 *Advanced Video Coding* (AVC), foi criado o grupo *Joint Collaborative Team on Video Coding* (JCT-VC) que reúne o grupo *Moving Picture Experts Group* (MPEG) do *International Standards Organization* (ISO) / *International Electrotechnical Commission* (IEC) e o grupo *Video Coding Experts Group* (VCEG) do *International Telecommunication Union - Telecommunication Standardization Sector* (ITU-T).

Em 2010, após a emissão formal de pedido de propostas (CfP – *Call for Proposals*) pelo grupo JCT-VC [15], na primeira reunião do grupo, foram analisadas 27 propostas. Algumas propostas demonstravam que era possível atingir a mesma qualidade que a norma H.264/MPEG-4 AVC, utilizando metade da taxa de transmissão, à custa de um aumento da complexidade computacional. Com a aprovação do nome *High Efficiency Video Coding* – HEVC (Codificação de Vídeo de Elevada Eficiência) para o projeto, foi iniciada a integração de técnicas das melhores propostas num único software de referência *Test*

Model under Consideration (TMuC) [16], resultando depois no *HEVC Test Model* (HM) [17].

A revisão de grande quantidade de contribuições nas diversas reuniões do JCT-VC, permitiram a otimização do HEVC, refletindo um melhor desempenho das sucessivas versões do software de referência, em termos de custo benefício entre eficiência de codificação e complexidade computacional.

Com o *Committee Draft* (*draft* completo da norma) anunciado em Fevereiro de 2012 e o *Draft International Standard* anunciado em Julho de 2012, em Janeiro de 2013 seguiu-se o *Final Draft International Standard*. Em Abril de 2013 o HEVC/H.265 foi aprovado como recomendação ITU-T H.265 ou ISO/IEC 23008-2, sendo publicado formalmente no *website* da ITU-T em Junho de 2013.

Várias extensões do HEVC continuam a ser desenvolvidas ativamente com o objetivo de que a norma suporte vários cenários de aplicação adicionais. Estas extensões incluem suporte de formatos de vídeo profissional com maior precisão, codificação escalável de vídeo e codificação de vídeo tridimensional (3D / estéreo / multivista).

Com o HEVC, tal como nas normas anteriores de codificação de vídeo, apenas a estrutura e sintaxe do fluxo binário (*bitstream*) é normalizada, bem como as restrições no fluxo binário e respetivo mapeamento para a obtenção de imagens decodificadas [5]. Qualquer decodificador que cumpra a norma irá produzir o mesmo resultado ao processar um fluxo binário normalizado. Deste modo, existe liberdade total para otimizar implementações em aplicações específicas.

3.2 Descrição Geral do Codificador

Tal como o seu predecessor, o H.264/MPEG-4 AVC, o HEVC utiliza uma estrutura de codificação híbrida baseada em blocos, combinação de predição espacial com compensação de movimento, transformada e codificação entrópica. Distingue-se, no entanto, na estrutura de particionamento flexível de imagem em *quadtree* (divisão recursiva de um bloco em quatro blocos quadrados), permitindo a utilização de blocos maiores, beneficiando a eficiência da codificação [4]. Com o HEVC foram também introduzidas melhorias ao nível da predição espacial e compensação de movimento, um novo filtro *in loop* e codificação entrópica, assim como a introdução de novas estruturas para processamento paralelo. A Figura 3.1 ilustra o diagrama geral do codificador. Nos parágrafos seguintes efetua-se a descrição geral das ferramentas utilizadas.

A codificação da imagem da sequência de vídeo é iniciada com a sua divisão em blocos de $N \times N$ amostras, denominados unidades de codificação em árvore (CTU – *Coding Tree Unit*), devido à sua estrutura em *quadtree*. A CTU é composta pelo bloco de codificação

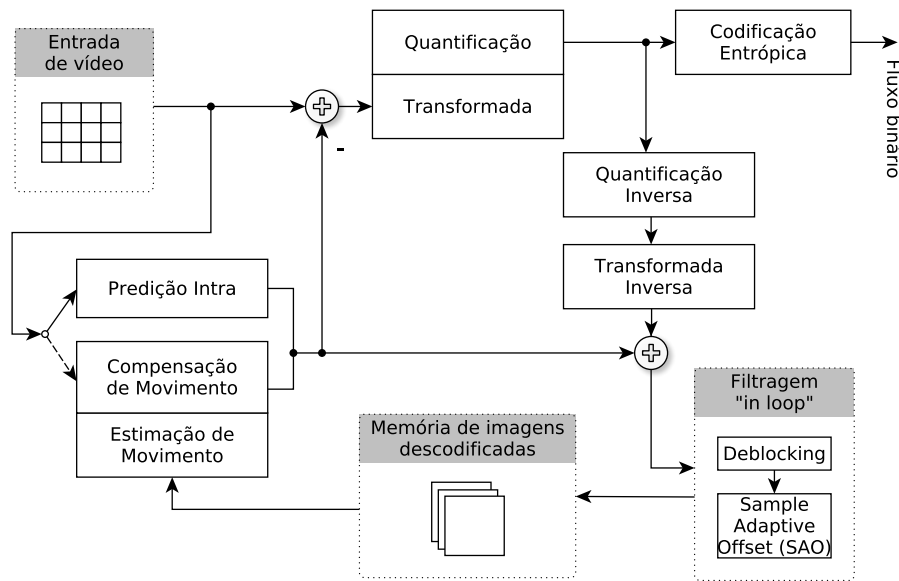


Figura 3.1: Diagrama geral de blocos do codificador HEVC.

em árvore (CTB – *Coding Tree Block*) da luminância que pode tomar os valores 64×64 , 32×32 ou 16×16 amostras e CTBs das correspondentes crominâncias. Na CTU os blocos de codificação (CB – *Coding Block*) de luminância e crominâncias formam a unidade de codificação (CU – *Coding Unit*) e são divididos de acordo com as características do sinal da região coberta pela respectiva CTB. Por sua vez, a CU contém unidades de predição (PU – *Prediction Unit*) com a predição de uma região da CU (Intra, Inter ou *Skipped*) e unidades de transformada (TU – *Transform Unit*). Uma CU *Skipped* é atribuída ao modo de predição Inter sem codificação das diferenças de vetores de movimento e informação do resíduo, enquanto que uma *non-Skipped* é atribuída ao modo de predição Intra ou Inter. A CU é sempre quadrada e pode ir de 8×8 (amostras de luminância) até ao tamanho da CTU, dependendo do conteúdo.

A primeira imagem de uma sequência de vídeo, bem como a primeira imagem de um ponto de acesso em vídeo de acesso aleatório, é codificada utilizando apenas o modo de predição Intra, ou seja, predição espacial a partir dos dados da própria imagem, não dependendo de outras imagens. Para as restantes imagens é utilizada a predição Inter para a maior parte dos blocos, que consiste na seleção do conteúdo em movimento relativo à imagem de referência e respetivos vetores de movimento (MV – *Motion Vector*). Estes dados são utilizados pelo codificador e decodificador para gerar dados de predição idênticos através da compensação de movimento (MC – *Motion Compensation*).

Cada CU pode conter uma ou mais unidades de predição (PU – *Prediction Unit*). A PU define uma região que partilha a mesma informação de predição, não está restringida à forma quadrada e os respetivos blocos de predição (PB – *Prediction Block*) podem ter um tamanho entre 4×4 e 64×64 . A unidade de transformada (TU – *Transform Unit*) na CU,

especificada por outra *quadtree*, define uma região que partilha a mesma transformada. O tamanho e a forma da TU depende da PU.

A Figura 3.2 mostra um exemplo de particionamento da imagem. Verifica-se deste modo, que ao contrário do que acontecia nas normas anteriores, o HEVC utiliza uma estrutura de partição *quadtree* de codificação de bloco flexível que permite o uso eficiente de blocos grandes e de múltiplos tamanhos de codificação, predição e transformada.

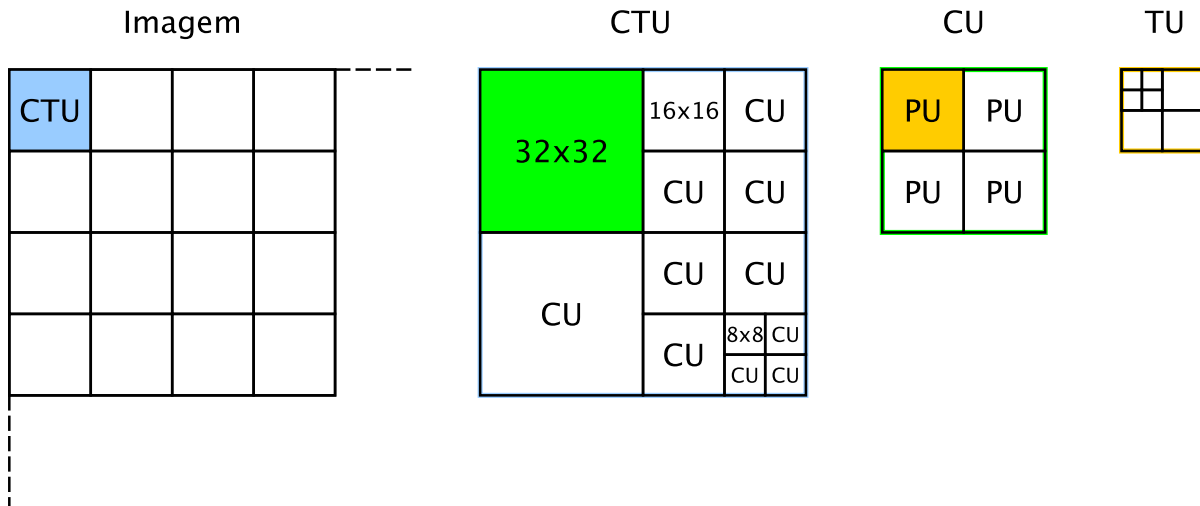


Figura 3.2: Exemplo simplificado de particionamento de uma imagem em CTU, CUs, PUs e TUs.

A estrutura *Slice* atua como uma unidade de empacotamento de vídeo codificado, tendo em vista a sua transmissão. É constituída pelo cabeçalho da *Slice* seguida das CUs. As *Slices* são independentemente descodificáveis, não sendo por isso efetuada predição para além do limite da *Slice*, sendo a codificação entrópica reiniciada em cada *Slice*. As *Slices* entrópicas diferem das anteriores pela possibilidade de existir predição entre elas. Se forem usados ambos os tipos de *Slices*, uma *Slice* pode conter várias *Slices* entrópicas, enquanto que uma *Slice* entrópica não pode incluir outras *Slices*.

A estrutura *Tile* foi adicionada ao HEVC tendo em vista o processamento paralelo. Possui maior flexibilidade que as *Slices*, mas menor complexidade. O *Tile* é especificado como uma unidade com um número inteiro de CTUs, definido por fronteiras verticais e horizontais, dividindo a imagem em regiões retangulares. As CTUs são processadas sequencialmente (*raster scan order*) em cada *Tile*, tal como os próprios *Tiles* o são na imagem.

De modo a ultrapassar a desvantagem, em termos de desempenho, devido à não existência de predição entre *Slices* ou *Tiles* e da reinicialização da codificação entrópica, foi adicionada a ferramenta de processamento paralelo *Wavefront* (WPP – *Wavefront Parallel Processing*). Neste caso, assim que é efetuado o processamento de dois CTUs numa linha (de CTUs), a descodificação da linha de CTUs seguinte poderá ser iniciada em paralelo,

usando informação já codificada.

Tal como mencionado anteriormente, a primeira imagem de uma sequência de vídeo (ou a primeira imagem de cada ponto de acesso aleatório na sequência) é codificada usando apenas predição de dados espaciais dessa mesma imagem, não dependendo de outras. A codificação Intra do HEVC tem um total de 35 modos de predição: planar, DC e 33 direções angulares.

A predição Intra utiliza amostras dos limites da divisão de PUs adjacentes, previamente decodificadas. A predição direcional para PUs quadradas de 4×4 até 32×32 amostras pode ter 33 direções diferentes, contrastando com as 8 direções do H.264/MPEG-4 AVC. É preciso, no entanto ter em conta que o número de direções de predição Intra depende do tamanho da CU. Para CUs 8×8 , 16×16 e 32×32 existem os 33 modos de predição angular, 16 modos para CUs 4×4 e 4 para CUs 64×64 (LCU) [18]. Na Figura 3.3 são ilustrados as direções angulares da predição Intra Angular com a dependência do tamanho da CU.

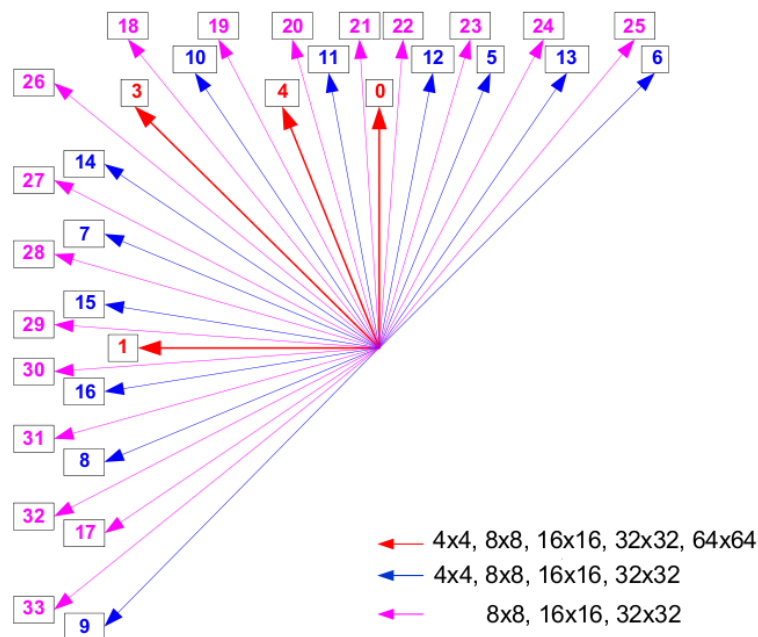


Figura 3.3: Direções angulares da predição Intra Angular de acordo com o tamanho da CU (adaptado de [18]).

Pode ser utilizada a predição planar no caso de superfícies com uma amplitude com um declive horizontal e vertical derivado dos limites da divisão de PUs adjacentes. No caso de uma superfície plana com um valor correspondente à média do valor do limite da divisão é utilizada a predição DC. Para a crominância os modos de predição horizontal, vertical, planar e DC, podem ser explicitamente sinalizados, ou sinalizados como sendo o mesmo modo de predição do bloco correspondente da luminância [5].

Um CB de tamanho $M \times M$ pode ter um PB particionado em $2N \times 2N$ ou $N \times N$, onde

$N=M/2$. Ou seja, com a partição $2N \times 2N$ o CB não é dividido, com $N \times N$ é dividido em quatro partes iguais. No HEVC apenas é permitido utilizar partições de predição $N \times N$, quando o tamanho do CB atual é igual ao tamanho mínimo de uma CU. Isto sucede devido à possibilidade de representar as mesmas regiões especificadas por quatro PBs usando quatro CBs de menor tamanho, quando o tamanho do CB atual é maior que o tamanho mínimo de uma CU [5].

A codificação Intra é suportada por todos os tipos de *Slice*. Dependendo do tipo de *Slice*, cada CB pode ser codificado por um de vários tipos de codificação Intra disponíveis: Intra Angular, Intra Planar e Intra DC.

A predição Intra Angular tem como objetivo atuar em regiões com contornos direcionais acentuados. As 33 direções diferentes de predição angular estão distribuídas de modo a fornecer uma cobertura mais densa para ângulos próximos da horizontal e vertical e mais espaçadas para ângulos próximos da diagonal, refletindo a prevalência estatística observada dos ângulos e melhorando a eficácia do processamento do sinal de predição [5]. O processo de predição dos modos Intra Angular é consistente com todos os tamanhos de bloco e direções de predição, ao contrário do H.264/MPEG-4 AVC que utiliza métodos diferentes para os tamanhos de bloco que suporta [5].

Tal como no H.264/MPEG-4 AVC, a predição Intra DC utiliza um valor médio das amostras de referência para a predição. A predição Intra Planar usa valores médios de duas predições lineares de quatro cantos de amostras de referência, para prevenir descontinuidades ao longo dos limites dos blocos [5].

A predição Inter é utilizada para codificar blocos de uma imagem usando outras previamente codificadas como referência. Este tipo de predição tira assim vantagem da similaridade de cada imagem com as suas imagens vizinhas temporais, explorando essa característica [19].

O HEVC suporta vários formatos de partição de blocos de predição PB. $2N \times 2N$ para um bloco de codificação CB não dividido, $2N \times N$ para PB dividido horizontalmente em duas partes iguais, $N \times 2N$ para um PB dividido em duas partes iguais verticalmente e $N \times N$ quando o CB, com o menor tamanho permitido, é dividido em quatro PBs com o mesmo tamanho [5]. Suporta ainda quatro tipos de partição assimétrica de movimento (AMP – Asymmetric Motion Partition) para divisão do CB em dois PBs de tamanho diferente: $2N \times nU$, $2N \times nD$, $nL \times 2N$ e $nR \times 2N$, onde “n” identifica a partição mais pequena e “U”, “D”, “L”, “R” representa cima, baixo, esquerda e direita, respetivamente [20]. A partição AMP melhora a eficiência de codificação, pois permite que os PBs se adaptem melhor à forma do objeto na imagem, sem necessidade de divisões adicionais [21]. Na Figura 3.4 pode-se observar os vários tipos de partição suportados na predição do HEVC.

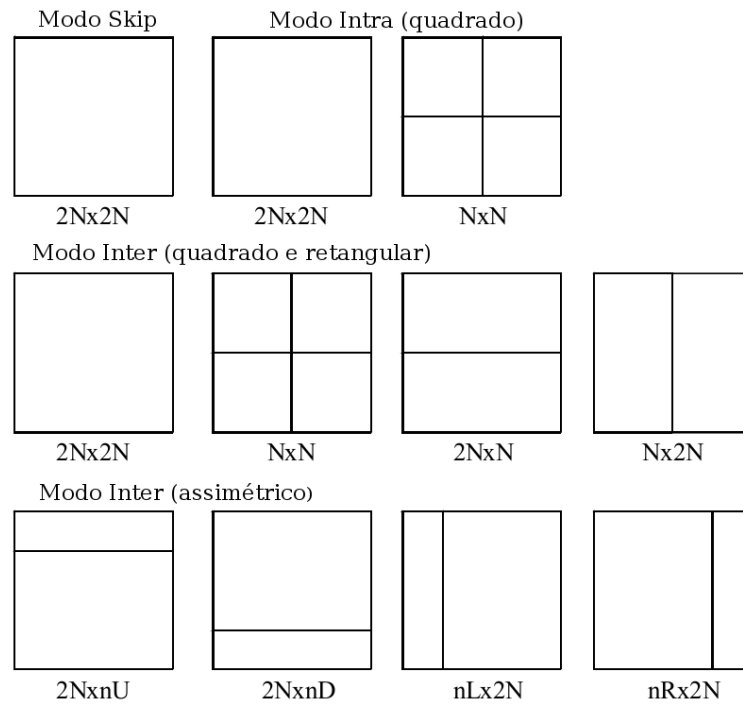


Figura 3.4: Tipos de partição suportados na predição do HEVC (adaptado de [20]).

As amostras de um bloco de predição PB para um bloco Inter codificado CB são obtidas a partir da região do bloco correspondente na imagem de referência numa posição deslocada pelas componentes horizontal e vertical do vetor de movimento. Para vetores de movimento de valor não inteiro, a interpolação de amostra fracionada é utilizada para gerar as amostras de predição para posições não inteiras de amostra [5].

O HEVC, assim como H.264/MPEG-4 AVC, suporta vetores de movimento com unidades de um quarto de distância entre as amostras de luminância. Nas amostras de crominância, a precisão do vetor de movimento é determinada de acordo com o formato da amostragem da croma. Para 4:2:0, a amostragem resulta em unidades de um oitavo da distância em amostras de croma [5].

A utilização de um processo consistente de interpolação separável para gerar todas as posições fracionadas sem operações intermédias com arredondamento permite ao HEVC melhorar a precisão e simplifica a arquitetura da interpolação fracionada de amostra [5]. Ao contrário do que acontece com o H.264/MPEG-4 AVC, no HEVC os valores de um quarto de pixel são calculados a partir do valor inteiro das amostras de luminância com um filtro maior, em vez da utilização de uma interpolação bilinear nos valores vizinhos de meio pixel e pixel inteiro [19]. A precisão é melhorada através da utilização de filtros de maior dimensão, de 7 ou 8 coeficientes, relativamente aos 6 coeficientes do H.264/MPEG-4 AVC [5].

Os vetores de movimento (MV – *Motion Vector*) podem ser preditos espacialmente, temporalmente ou através da técnica *Motion Merge*. Para cada PB codificado no modo

Intra, o codificador pode escolher entre a codificação explícita dos parâmetros de movimento, o modo *Motion Merge* ou utilizar o modo *Skipped* melhorado [19].

O modo *Motion Merge* cria uma lista de PBs vizinhos, espaciais ou temporais, previamente codificados, candidatos ao PB a ser codificado. O codificador sinaliza qual o candidato a ser utilizado e a informação de movimento é copiada do candidato selecionado para o PB atual. Deste modo, esta técnica evita a necessidade de codificar o vetor de movimento para o PB, sendo apenas necessário codificar o índice do candidato da lista [19].

No modo *Skipped* o HEVC também codifica o índice do candidato da lista *Motion Merge* e os parâmetros de movimento para o PB atual são copiados do candidato selecionado. Assim, torna-se possível codificar, com poucos bits, zonas da imagem que sofram poucas alterações entre imagens ou que tenham um movimento constante [19].

No caso da codificação explícita de parâmetros de movimento, a diferença do vetor de movimento é codificada utilizando um preditor do vetor de movimento, escolhido entre múltiplos candidatos preditores. A diferença entre o preditor, o vetor de movimento atual e o índice do candidato são transmitidos ao decodificador [5].

O sinal residual da predição Intra ou Inter, ou seja, a diferença entre o bloco original e a sua predição, é transformado por uma transformada espacial linear. Os seus coeficientes são depois escalados, quantificados, codificados entrópicamente e transmitidos em conjunto com a informação de predição [5].

O bloco residual é particionado em múltiplos blocos quadrados denominados blocos de transformada (TB – *Transform Block*), com tamanhos 4×4 , 8×8 , 16×16 e 32×32 . A transformada bidimensional é realizada aplicando a transformada de uma dimensão em ambas as direções horizontal e vertical. Os elementos das matrizes de transformada foram obtidos aproximando funções base escaladas da transformada discreta de cosseno (DCT – *Discrete Cosine Transform*) [5]. Apenas uma matriz de valores inteiros é especificada para o tamanho de 32 pontos. Para simplificar, para outros tamanhos são utilizadas as versões subamostradas dessa matriz.

Em blocos codificados com alguns modos de predição Intra, para a transformada de tamanho 4×4 é utilizada uma transformada alternativa, a transformada discreta do seno (DST – *Discrete Sine Transform*). Ao utilizar a predição Intra, os *pixels* próximos dos *pixels* dos limites do topo ou da esquerda de um bloco usados para a predição são habitualmente preditos com maior exatidão que os *pixels* mais afastados. Assim, o resíduo tem tendência a ser mais elevado para *pixels* afastados dos limites do bloco. A DST codificará melhor este resíduo, porque as funções base iniciam com uma amplitude baixa e depois aumentam para frequências mais elevadas, enquanto que as funções base da DCT iniciam com uma amplitude elevada e depois diminuem [19] para as frequências

mais elevadas.

Depois da transformada, a codificação entrópica codifica todos os elementos de sintaxe e coeficientes de transformada quantificados. A codificação entrópica usa uma versão melhorada, em relação ao H.264/MPEG-4 AVC, da codificação aritmética binária com adaptação ao contexto (CABAC – *Context Adaptive Binary Arithmetic Coding*). Apresenta melhorias na complexidade (especialmente para processamento paralelo), no desempenho de compressão e na redução de necessidade de memória para o seu contexto [5]. A codificação entrópica é aplicada aos símbolos gerados e coeficientes da transformada quantificados, dando origem ao fluxo binário (*bit stream*).

O processo de filtragem *in loop*, destina-se a obter melhor eficiência de codificação e qualidade visual. Consiste na atuação de duas ferramentas de filtragem: *Deblocking* (DBF), seguido de *Sample Adaptive Offset* (SAO) após a reconstrução da imagem. O DBF é semelhante ao utilizado no H.264/MPEG-4 AVC, tendo como tarefa reduzir os artefactos produzidos pelas operações efetuadas sobre blocos da imagem, sendo aplicado às amostras nas fronteiras dos blocos. O SAO é aplicado de forma adaptativa a todas as amostras que satisfazem determinadas condições (por exemplo, com base no gradiente) [5].

Após a reconstrução da imagem, esta é guardada na memória (*buffer*) de imagens descodificadas, de modo a poder ser utilizada para codificar a próxima imagem.

O HEVC permite também vários modos de codificação especial, nomeadamente, o modo de codificação sem perdas (*lossless*). A compressão sem perdas é útil quando se pretende a preservação rigorosa da informação contida nas imagens. Algumas aplicações, tais como, imagens médicas, imagens de obras de arte, análise de imagem, requerem a utilização de compressão sem perdas de modo a evitar qualquer distorção nas imagens reconstruídas [22].

Neste modo, algumas ferramentas de codificação no HEVC não são utilizadas, como por exemplo, a transformada, quantificação e outros processos que afetam a imagem descodificada, nomeadamente, o SAO, os filtros *in loop* e *Deblocking*. O sinal residual da predição Intra ou Inter alimenta diretamente o codificador entrópico [5]. Deste modo, é possível efetuar uma reconstrução sem perdas, sem necessidade de utilizar ferramentas de codificação adicionais.

Capítulo 4

Codificação Compatível 3D

4.1 Codificação assimétrica

Esta dissertação aborda o tema da transmissão simultânea de vídeo estéreo assimétrico, onde uma vista base é codificada com resolução original e a vista auxiliar é subamostrada na horizontal e/ou vertical para metade do seu valor antes da compressão. Após a decodificação, a vista auxiliar é sobreamostrada para a resolução original. A Figura 4.1 ilustra o diagrama de blocos de um sistema genérico de transmissão simultânea de vídeo estéreo assimétrico, como o considerado.

Neste trabalho é proposta uma técnica de melhoria da qualidade da vista auxiliar decodificada, aplicada ao decodificador assimétrico descrito. Esta técnica visa compensar a perda de qualidade da vista auxiliar que resulta da respetiva subamostragem e posterior sobreamostragem, após a decodificação. Os detalhes (altas frequências) que se perdem da vista auxiliar são estimados com base na informação disponível na vista de resolução original. Esta informação de detalhe é determinada através da diferença entre a vista base decodificada e uma versão desta imagem obtida pelo mesmo processo de sub

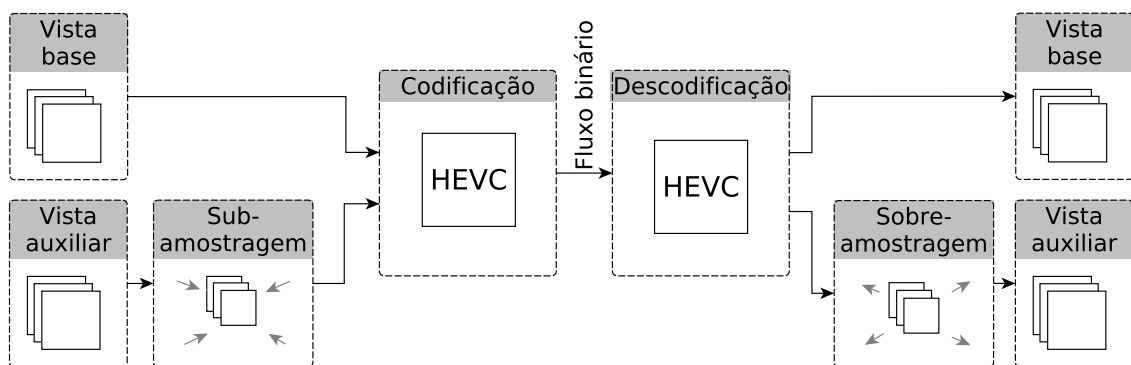


Figura 4.1: Diagrama geral de transmissão *simulcast* assimétrica.

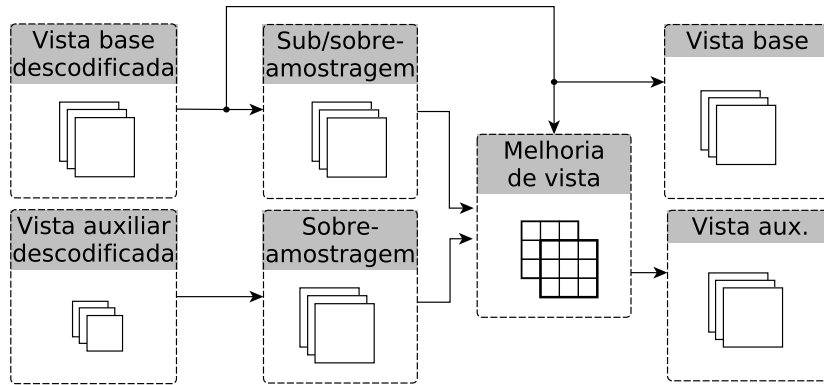


Figura 4.2: Arquitetura geral do decodificador.

e sobreamostragem utilizado na codificação da vista auxiliar. O diagrama de blocos do decodificador assimétrico que usa este princípio é ilustrado na Figura 4.2.

Os detalhes da vista base não podem ser adicionados diretamente à vista auxiliar sobreamostrada, devido à disparidade existente entre o conteúdo das duas imagens. Para efetuar a adição, torna-se assim necessário utilizar um método de compensação de disparidade. As abordagens tradicionais a este problema utilizam normalmente correspondência de blocos (*block matching*) [23], como descrito na secção 4.2. De modo a colmatar os problemas relacionados com as imprecisões destes métodos, neste trabalho é proposto um novo procedimento, que usa um mapeamento usando uma malha triangular (*triangular mesh*), que é descrito na secção 4.3.

4.2 Método baseado em pesquisa de bloco

Tal como foi descrito na secção anterior, a imagem de detalhes é obtida subtraindo à vista base descodificada, uma versão desta imagem que foi subamostrada e posteriormente sobreamostrada. Segue-se um passo de correspondência de blocos semelhantes entre as versões descodificadas da vista auxiliar sobreamostrada e da vista base. Este procedimento permite que os detalhes determinados no primeiro passo sejam adicionados à vista auxiliar, bloco a bloco, nos sítios apropriados.

Para efetuar a pesquisa de bloco semelhante, a vista auxiliar é dividida em blocos quadrados de tamanho fixo. Cada bloco é usado como referência para procurar um bloco semelhante na vista base numa janela de pesquisa, obtendo-se assim a respetiva estimação de disparidade. A janela de pesquisa é geralmente de secção retangular na horizontal, já que as sequências de vídeo 3D estéreo são tipicamente adquiridas usando câmaras com eixos óticos alinhados na horizontal. Na Figura 4.3 pode-se observar o esquema de funcionamento deste processo. Nos casos em que existe apenas disparidade na horizontal, tipicamente o conteúdo da vista esquerda vai estar localizado ligeiramente à direita em

relação ao conteúdo da vista direita.

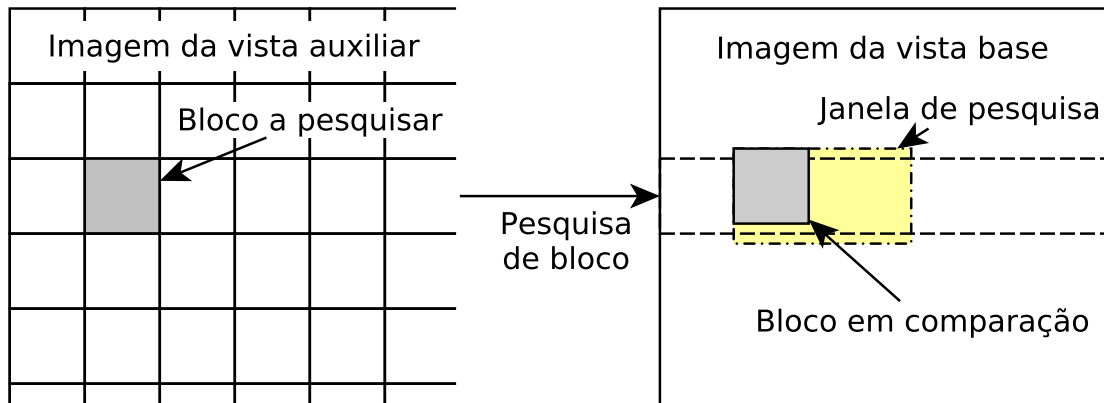


Figura 4.3: Pesquisa de bloco semelhante na vista auxiliar.

A estimação de disparidade é efetuada através de um algoritmo que determina as condições que minimizam a distorção entre blocos de imagens. Neste caso é utilizado o valor de soma absoluta das diferenças (SAD – *Sum of Absolute Differences*), obtida entre o bloco de referência e os blocos existentes na janela de pesquisa. A SAD pode ser expressa através da equação:

$$\sum_{(i,j) \in W} |I_1(i, j) - I_2(x + i, y + j)|, \quad (4.1)$$

onde I_1 e I_2 correspondem às matrizes com os valores dos *pixels* da componente de luminância de cada bloco. No caso do segundo bloco, as coordenadas de cada pixel estão deslocadas de um vetor com coordenadas pertencentes à janela de pesquisa, W .

Os valores obtidos para a disparidade horizontal e vertical, (x, y) , que minimizam a SAD para um dado bloco são depois usados como deslocamento. Este deslocamento é adicionado à posição do bloco com os detalhes da vista base, de modo a obter o bloco a somar à vista auxiliar. Este processo é repetido para todos os blocos, definindo assim uma imagem de detalhes completa e na resolução original da sequência, que será depois adicionada à vista auxiliar decodificada e sobreamostrada.

4.3 Método baseado em malha triangular

O método proposto neste trabalho para melhorar a vista subamostrada, baseia-se no mapeamento de uma malha (*mesh*) triangular criada sobre a imagem de contornos (*edges*) da vista auxiliar. Através da adaptação do conteúdo de uma malha triangular na imagem, o algoritmo proposto reduz os efeitos de bloco tipicamente observados no método anterior [23], melhorando assim a qualidade da imagem reconstruída. A Figura 4.4 ilustra o

diagrama de blocos do método proposto.

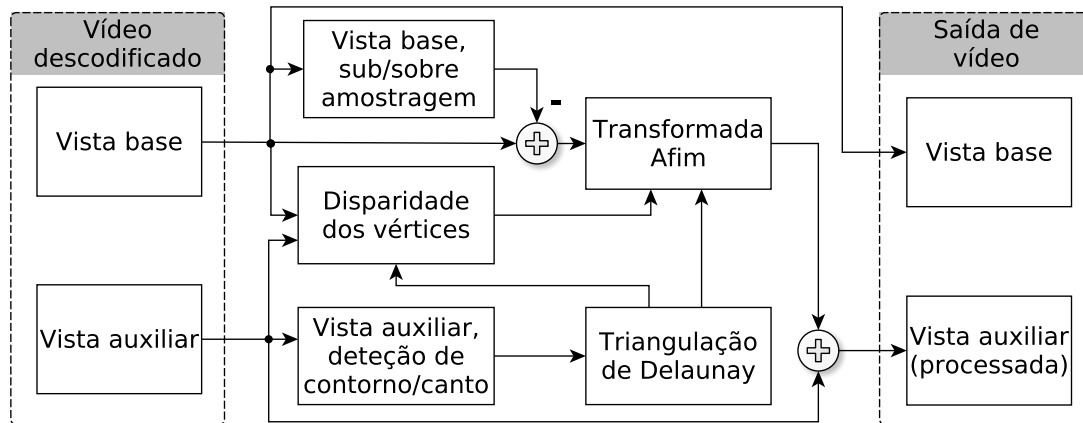


Figura 4.4: Diagrama do método da malha triangular.

Para a aplicação de uma malha triangular numa imagem estéreo que possibilite a adaptação do conteúdo de uma das vistas à outra, será necessário encontrar pontos nessa imagem localizados sobre as fronteiras dos objetos, de modo a facilitar o processo de adaptação. A utilização de um algoritmo de detecção de contornos, como por exemplo o algoritmo Canny [24], permite obter as fronteiras. Com a aplicação de um algoritmo de detecção de cantos (*corners*) na imagem de contornos, como por exemplo o algoritmo Shi-Tomasi [25], são obtidos pontos dispersos pelos contornos, localizados sobretudo em interseções de dois contornos ou onde existam duas direções de contornos dominantes na vizinhança do ponto, úteis para criar os vértices da malha triangular.

Assim, inicialmente é aplicada a detecção de contornos com o algoritmo Canny na vista auxiliar, seguindo-se a detecção de cantos (*corners*) com o algoritmo Shi-Tomasi, de modo a obter os pontos correspondentes aos cantos do conteúdo da imagem. A seleção específica destes algoritmos pode-se considerar arbitrária, sendo necessário a realização de testes adicionais de modo a determinar quais os algoritmos que apresentam melhores resultados.

A detecção de contornos permite reconhecer fronteiras que delimitam diferentes características no conteúdo da imagem. Neste caso, como a variação de disparidade das imagens depende da profundidade, as características com maior interesse em evidenciar são as descontinuidades na profundidade.

O detetor de contornos Canny consiste num algoritmo de múltiplos passos:

- Redução de ruído da imagem através de um filtro Gaussiano, resultando numa imagem ligeiramente desfocada, para evitar detecção de falsos contornos;
- Localização do gradiente de intensidade da imagem, utilizando quatro filtros para detetar contornos horizontais, verticais e diagonais na imagem do passo anterior;

- Pesquisa para verificar a direção do contorno com base na direção do gradiente. Neste passo é obtido um conjunto de pontos dos contornos na forma de imagem binária;
- Aplicação de níveis (*thresholds*) com histerese para marcar e traçar pontos de contornos. A utilização dos níveis alto (*high*) e baixo (*low*) permite seguir secções ténues de uma linha, descartando *pixels* de ruído que não fazem parte dessa linha.

O algoritmo de deteção de cantos Shi-Tomasi é baseado no algoritmo Harris *Corner Detector* [26], com uma modificação no critério de seleção. Com o Harris *Corner Detector* é calculado um valor (pontuação) \mathcal{R} para cada pixel a partir de dois valores próprios λ_1 e λ_2 de duas regiões, através da equação:

$$\mathcal{R} = \det(\lambda_1 \lambda_2) - k(\lambda_1 + \lambda_2)^2. \quad (4.2)$$

Esta função permite determinar a existência de canto em regiões com uma variação elevada ao efetuar um pequeno deslocamento em qualquer direção. Se o valor obtido for superior a um determinado nível predefinido, o pixel é marcado como canto.

O algoritmo Shi-Tomasi propõe a utilização dos valores próprios para verificar se o pixel deve ser considerado um canto ou não, de acordo com a equação:

$$\mathcal{R} = \min(\lambda_1, \lambda_2) \quad (4.3)$$

Os autores demonstraram que este critério para atribuição de pontuação obtém bons resultados.

A triangulação de Delaunay [27] é aplicada à imagem da vista auxiliar utilizando como vértices os pontos obtidos no passo anterior. Com este processo obtém-se um conjunto de pontos num plano, que não se encontram localizados dentro do círculo circunscrito por uma circunferência que passa por todos os vértices de qualquer triângulo. A triangulação de Delaunay maximiza o menor ângulo de todos os triângulos da triangulação, evitando assim o aparecimento de triângulos com ângulos internos muito pequenos. A Figura 4.5 ilustra um exemplo onde cada triângulo satisfaz este critério.

Esta operação resulta numa tesselação não regular com triângulos que limita zonas de textura diferente, como se pode observar no exemplo da Figura 4.6.

No método proposto, a imagem de detalhes a adicionar à vista auxiliar resulta do mapeamento da imagem de detalhes da vista base, obtido através de aplicação de uma deformação da *mesh* definida nos passos anteriores. Para este efeito aplica-se um procedimento de estimação de disparidade entre a vista auxiliar e a vista base, aplicada a pequenos blocos (*kernels*), centrados em cada um dos vértices dos triângulos determina-

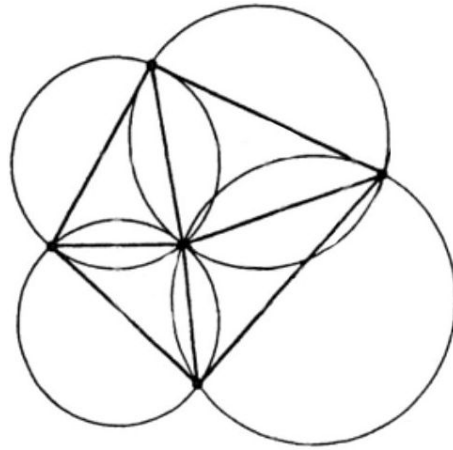
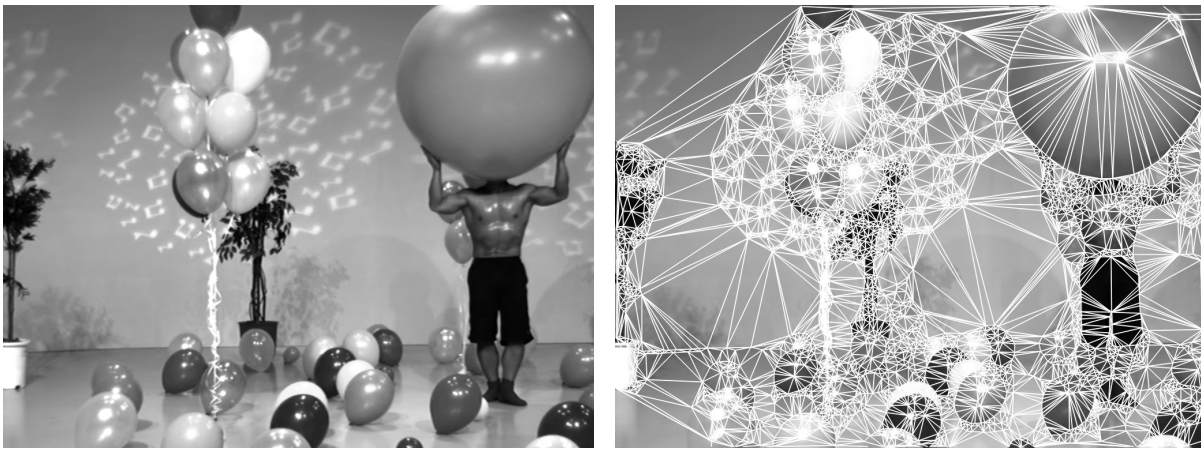


Figura 4.5: Exemplo de triangulação de Delaunay (adaptado de [28]).



(a) *Balloons* (câmara 5, Imagem 1).

(b) Após a aplicação da triangulação.

Figura 4.6: Triangulação de Delaunay aplicada à Imagem 1 da sequência *Balloons*.

dos para a vista auxiliar. Tal como no caso da correspondência de blocos é utilizada uma janela de pesquisa, definida na imagem da vista base. A tesselação não regular da vista base é formada deslocando os vértices dos triângulos da vista auxiliar, de acordo com a disparidade obtida para cada um deles. Na Figura 4.7 pode-se observar um possível exemplo de um triângulo obtido na vista auxiliar, assim como as janelas de pesquisa de cada vértice na vista base e o deslocamento efetuado após a obtenção da disparidade.

A transformação de um triângulo da imagem auxiliar no triângulo correspondente da imagem base é efetuada através da aplicação de uma transformada Afim [29, 30]. Trata-se de uma transformação que preserva a colinearidade, ou seja, todos os pontos inicialmente numa linha irão continuar numa linha após a transformação, preservando também as razões entre distâncias. De uma maneira geral, a transformada Afim é uma composição de rotações, translações, dilatações e deformações do tipo *shear* (transformações onde todos os pontos de uma determinada linha se mantêm fixos, enquanto que outros pontos são

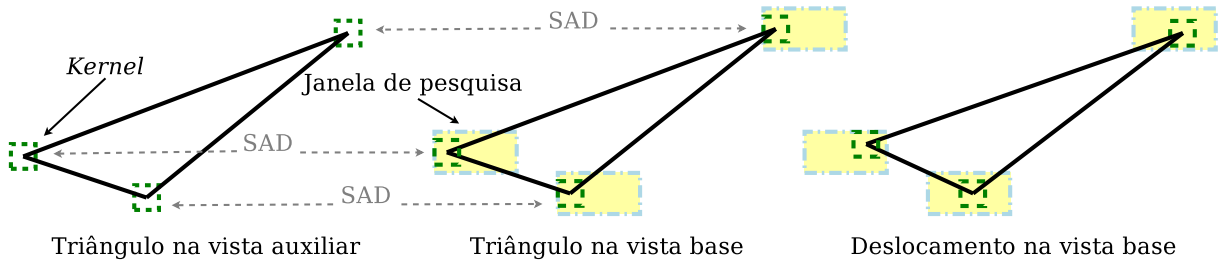


Figura 4.7: Triângulo na vista auxiliar, aplicado na vista base e deslocado de acordo com a disparidade.

deslocados paralelamente a essa linha de forma proporcional à sua distância perpendicular à linha). Não preserva necessariamente ângulos ou distâncias.

A transformada Afim pode representar a relação entre duas imagens, por exemplo através de uma matriz de multiplicação seguida de adição de um vetor. Considerando

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix},$$

$$B = \begin{bmatrix} b_{00} \\ b_{10} \end{bmatrix},$$

$$M = \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} & b_{00} \\ a_{01} & a_{11} & b_{10} \end{bmatrix},$$

a transformada de um vetor bidimensional, $X_i = [x, y]^T$ utilizando A e B , pode-se obter de duas formas:

$$X_f = A \cdot \begin{bmatrix} x \\ y \end{bmatrix} + B,$$

ou

$$X_f = M \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T,$$

obtendo-se assim:

$$X_f = \begin{bmatrix} a_{00} \cdot x + a_{01} \cdot y + b_{00} \\ a_{01} \cdot x + a_{11} \cdot y + b_{10} \end{bmatrix}.$$

A intensidade de um pixel localizado em X_i é mapeado para uma nova posição X_f . Sabendo as coordenadas iniciais e finais de um vértice, calcula-se M de modo a aplicar essa relação a todos os pixels da região que se pretende transformar.

A transformada pode assim produzir qualquer paralelogramo a partir de um retângulo.

Mapeia pontos de uma região de uma imagem para uma região diferente, muitas vezes efetuando interpolação subpixel nesse processo. Aplicando a malha triangular à imagem de detalhes (alta frequência) e utilizando a informação da disparidade de cada vértice, com a transformada Afim é possível efetuar o mapeamento da imagem de detalhes, aplicando a relação obtida ao conteúdo correspondente a cada triângulo.

Cada triângulo da imagem de detalhes é adaptado à imagem auxiliar através do processo descrito anteriormente, de modo que a nova imagem de detalhes possa ser adicionada à imagem auxiliar. A imagem resultante, com os detalhes da vista base deslocados e transformados, é adicionada à vista auxiliar do mesmo modo que no método de pesquisa de blocos, usando o processo descrito na secção seguinte.

4.4 Adição da imagem de detalhes

O método baseado na malha triangular e o método baseado na pesquisa de blocos processam a imagem de detalhes da vista base, preparando-a para ser adicionada à imagem da vista auxiliar. No entanto, devido às diferenças entre a vista base e a auxiliar, a adição dos detalhes da vista base nem sempre é benéfica para a vista auxiliar. Deste modo, foi estudado um algoritmo que seleciona quais as regiões da imagem em que a adição de detalhes contribui para melhorar a qualidade da vista auxiliar sobreamostrada. A decisão de adicionar um determinado bloco de detalhes à vista auxiliar depende de uma máscara que sinaliza as zonas (geralmente não homogêneas) que possam beneficiar desta adição. O diagrama de blocos da Figura 4.8 ilustra o método de adição seletiva de detalhes.

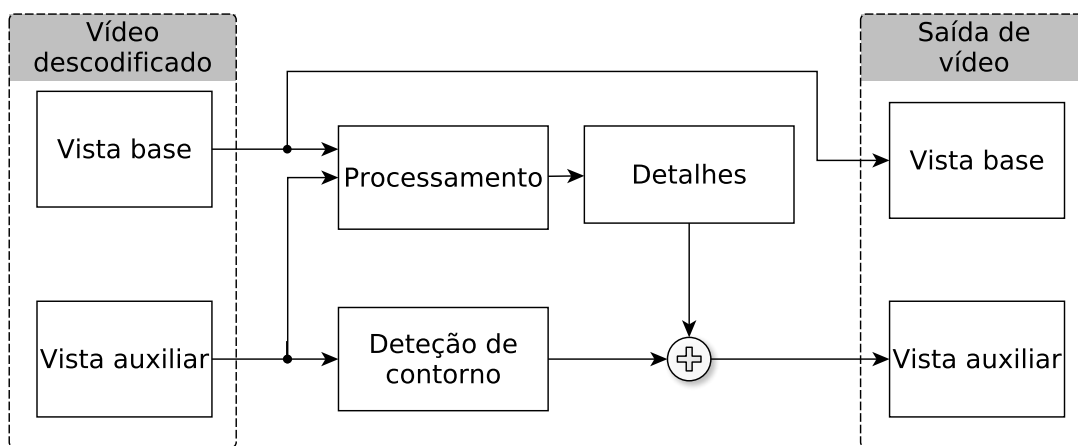


Figura 4.8: Método de adição seletiva de detalhes.

O desenvolvimento do método proposto teve como base um teste preliminar que avaliou em que regiões a utilização do detalhe de alta frequência contribui para a melhoria da imagem auxiliar. Para este teste foi realizada a partição da imagem original, da imagem auxiliar e da imagem de detalhes, em blocos de tamanho fixo, 8×8 , sendo os blocos de

detalhes adicionados condicionalmente para cada uma destas regiões. Um determinado bloco da imagem de detalhes da vista base é utilizado apenas quando a respetiva adição ao bloco correspondente da vista auxiliar resultar numa maior qualidade objetiva (PSNR – *Peak Signal-to-Noise Ratio*). Se o resultado da adição diminuir a qualidade objetiva, então o bloco de detalhes não é adicionado ao respetivo bloco da imagem auxiliar.

Por exemplo, para o caso da Figura 4.9 (a), 30% dos blocos revelaram poder ser melhorados com a soma dos detalhes da vista base. Estes blocos são assinalados a branco na Figura 4.9 (b). Já para o exemplo da Figura 4.9 (c), a percentagem dos blocos para os quais deve ser utilizado o passo de melhoria sobe para 48%, representados na Figura 4.9 (d).

(a) *Kendo* (Câmara 5, Imagem 1).

(b) Localização de 30% de blocos adicionados.

(c) *Book Arrival* (Câmara 10, Imagem 1).

(d) Localização de 48% de blocos adicionados.

Figura 4.9: Exemplos do modo de decisão utilizando as imagens originais.

Como pode ser observado nos exemplos da Figura 4.9, a maior parte dos blocos escolhidos situa-se nas zonas não homogêneas da imagem, por exemplo, nas fronteiras dos objetos. No entanto, nestas zonas também pode ser difícil recuperar informação, já que

são zonas mais sujeitas a oclusões. Também pode ser observado na Figura 4.9 que existem alguns detalhes em zonas homogêneas que sendo adicionados contribuiriam para melhorar a vista auxiliar. Tal como esperado, quando vértices dos triângulos ou blocos no método de pesquisa de blocos surgem em zonas mais uniformes, torna-se mais difícil estimar a disparidade, devido à homogeneidade dos valores dos *pixels*.

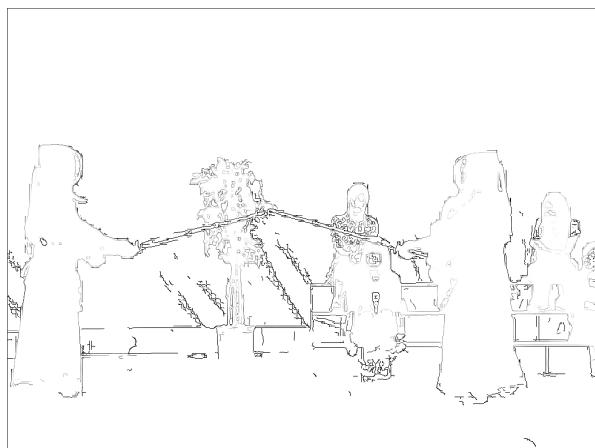
Este teste foi realizado utilizando a versão original da vista auxiliar. Num cenário real, esta vista não está, obviamente, disponível no decodificador. Deste modo foi desenvolvido um algoritmo para decidir quais as regiões que deveriam ser selecionadas para adicionar detalhes da vista base.

Com base nas observações realizadas no teste descrito anteriormente, o critério definido para a utilização da informação de detalhe para um dado bloco foi a presença de um contorno nessa área da imagem. Deste modo, inicialmente é aplicado um algoritmo de detecção de contornos (Canny) à imagem da vista auxiliar. Na Figura 4.10 (a) e (c) mostram-se as imagens de contornos dos exemplos da Figura 4.9.

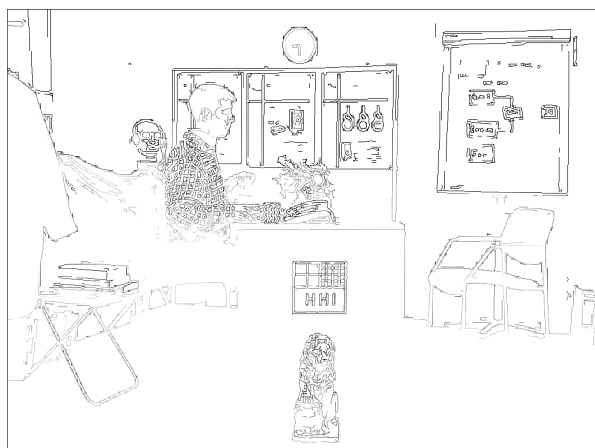
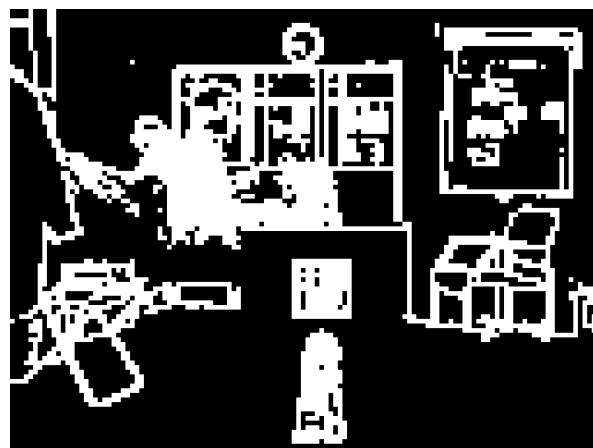
A imagem de contornos é segmentada com o mesmo padrão de blocos 8×8 definido anteriormente. Seguidamente são assinalados os blocos onde se verifica a presença de contornos de modo a criar a máscara que define as regiões para as quais deve ser adicionada a informação de detalhe. Assim, os detalhes são adicionados apenas nas zonas delimitadas pela máscara, que correspondem às zonas que envolvem os contornos.

Para os casos já analisados, a aplicação do algoritmo proposto resulta na adição de informação de detalhe a 17% de blocos da imagem da Figura 4.9 (a) e a 27% dos blocos da imagem da Figura 4.9 (c). Estes blocos são representados nas Figuras 4.10 (b) e (d), respetivamente.

A adição da imagem de detalhes à imagem auxiliar sobreamostrada, de acordo com o algoritmo proposto, demonstrou ser favorável para a generalidade das situações. Os resultados dos testes de aplicação dos métodos descritos neste capítulo são analisados no Capítulo 5.

(a) Imagem de contornos (*Kendo*).

(b) Localização de 17% de blocos adicionados.

(c) Imagem de contornos (*Book Arrival*).

(d) Localização de 27% de blocos adicionados.

Figura 4.10: Exemplos com o modo de decisão utilizando a imagem de contornos.

Capítulo 5

Resultados Experimentais

A avaliação de desempenho taxa-distorção (RD – *Rate-Distortion*) foi efetuada através de testes experimentais utilizando sequências de teste de vídeo publicamente disponíveis. Previamente à avaliação de desempenho taxa-distorção dos métodos analisados, foram estabelecidos os respectivos parâmetros de processamento.

Os testes de simulação são realizados utilizando um conjunto de sequências de teste, com os mesmos parâmetros de configuração do codificador do modelo de teste (HM 13.0) da norma H.265/*High Efficiency Video Coding* (HEVC).

Dentro dos parâmetros específicos do método de pesquisa de blocos são considerados o tamanho de bloco a utilizar na estimação de disparidade e o tamanho da janela de pesquisa. Do mesmo modo, para o método baseado na malha triangular é analisado o tamanho do bloco centrado no vértice de cada triângulo, bem como o tamanho da janela de pesquisa. Outro parâmetro a ter em conta é a localização das margens da imagem que não fazem parte das vistas base e auxiliar. Estas zonas não devem ser processadas.

Em ambos os métodos, para o processo de pesquisa de blocos é necessário selecionar qual a vista que disponibiliza a imagem de referência com os blocos a pesquisar na outra vista. O objetivo é conhecer o deslocamento de determinado bloco da imagem, que pode ser obtido utilizando o bloco na vista auxiliar como referência a localizar numa janela de pesquisa da vista base, ou a situação inversa. Para selecionar a vista de referência torna-se necessário efetuar testes com ambas as configurações possíveis e verificar qual a mais favorável nos métodos em teste.

De seguida descrevem-se as sequências utilizadas e testes realizados para obter os parâmetros para cada método, seguindo-se a apresentação e discussão dos resultados.

5.1 Sequências de teste

As sequências de vídeo estéreo selecionadas para a aplicação dos métodos em comparação, têm 8 bits por amostra e foram codificadas com o software de referência (HM 13.0). A codificação das sequências foi efetuada com 64 imagens (*frames*) iniciais de cada par estéreo utilizando a configuração de acesso aleatório e perfil *Main* (MP – *Main Profile*), comparável em termos de eficiência de codificação ao perfil *High* (HP – *High Profile*) do H.264/MPEG-4 *Advanced Video Coding* (AVC) [31].

Foram utilizados os mesmos parâmetros de quantificação (QP – *Quantization Parameter*) usados nas condições comuns de teste do HEVC (*core experiments common test conditions*) [32]: 22, 27, 32 e 37. As sequências foram codificadas com as componentes de luminância (Y) e crominâncias (Cb,Cr), no entanto, apenas a luminância foi processada com os métodos em estudo, visto tratar-se da componente de maior relevância. A Tabela 5.1 lista as sequências de vídeo estéreo utilizadas nos testes.

Tabela 5.1: Sequências estéreo testadas.

Nome da sequência	Resolução	Câmara 1	Câmara 2
<i>GT Fly</i>	1920 × 1088	9	5
<i>PoznanHall 2</i>	1920 × 1088	7	6
<i>Pantomime</i>	1280 × 960	40	41
<i>Champagne Tower</i>	1280 × 960	41	42
<i>Balloons</i>	1024 × 768	4	5
<i>Kendo</i>	1024 × 768	4	5
<i>Book Arrival</i>	1024 × 768	11	10
<i>Ballroom</i>	640 × 480	4	5
<i>Exit</i>	640 × 480	4	5

A sequência de animação no exterior *GT Fly* [33] tem detalhe elevado, sem movimento de objetos mas com movimento da câmara. A imagem apresenta uma complexidade elevada de estrutura da profundidade e zonas significativas de conteúdo homogêneo que diminuem ao longo da sequência, devido ao movimento de câmara.

Na sequência *Poznan Hall 2* o conteúdo foi obtido em ambiente interior, tem movimento complexo de objetos, com reflexões e movimento de câmara. Apresenta um nível de detalhe médio e uma estrutura de profundidade de média complexidade. Nas condições de captura de imagem, todas as câmaras encontram-se alinhadas, com eixos óticos paralelos que distam entre si 13.75 cm e 110.0 cm entre os eixos óticos das câmaras mais afastadas [34].

Nas Figuras 5.1 (a) e (b) pode-se observar uma imagem de uma das vistas das sequências *GT Fly* e *Poznan Hall 2*, respetivamente.



(a) *GT Fly* (Câmara 5, Imagem 1). (b) *Poznan Hall 2* (Câmara 6, Imagem 1).

Figura 5.1: Sequências estéreo utilizadas nos testes, com resolução 1920×1088 .

A sequência *Champagne Tower* foi obtida em local interior do tipo estúdio, apresenta um detalhe elevado, tem movimento complexo de objetos, tais como, líquidos, reflexões e transparências. Nas condições de captura de imagem a matriz com as câmaras equidistantes encontra-se em posição fixa, com as câmaras distanciadas entre si 50 mm, convergindo para o centro da cena que está a uma distância de 8.2 m [35]. A sequência tem uma complexidade média em termos de estrutura da profundidade e um fundo estático e homogêneo. Na Figura 5.2 (a) observa-se um exemplo ilustrativo de uma imagem desta sequência.

A sequência *Pantomime* foi obtida com a mesma configuração de câmaras que a *Champagne Tower*, tendo por isso características semelhantes. Apresenta, no entanto, menor complexidade de movimento de objetos e profundidade. A Figura 5.2 (b) representa um exemplo de uma imagem desta sequência.



(a) *Champagne Tower* (Câmara 41, Imagem 1). (b) *Pantomime* (Câmara 40, Imagem 1).

Figura 5.2: Sequências estéreo utilizadas nos testes, com resolução 1280×960 .

Na sequência *Balloons* o conteúdo foi obtido em ambiente interior do tipo estúdio, com grande quantidade de objetos em movimento, reflexões e transparências. Parte signi-

ficativa do conteúdo de fundo é dinâmico com detalhe médio, existindo no entanto, algum conteúdo homogêneo. A câmara encontra-se em movimento horizontal. Esta sequência apresenta uma estrutura de profundidade com complexidade média. A Figura 5.3 (a) ilustra a Imagem 1 da câmara 4 desta sequência.

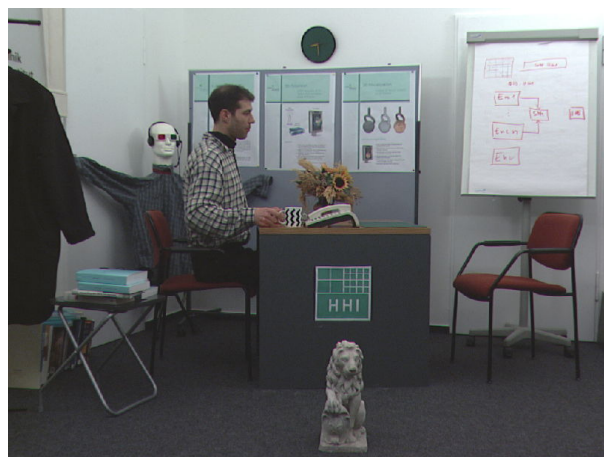
A sequência *Kendo* apresenta um ambiente interior do tipo estúdio, com movimento complexo de objetos, nomeadamente, fumo, reflexões e transparências. Parte significativa do conteúdo de fundo é estático e homogêneo. A câmara encontra-se em movimento horizontal. Esta sequência apresenta uma estrutura de profundidade com complexidade média. Na Figura 5.3 (b) pode-se observar um exemplo de uma imagem desta sequência. Ambas as sequências *Kendo* e *Balloons* foram obtidas com o mesmo dispositivo de captura descrito em [35].



(a) *Balloons* (Câmara 4, Imagem 1).



(b) *Kendo* (Câmara 4, Imagem 1).



(c) *Book Arrival* (Câmara 11, Imagem 1).

Figura 5.3: Sequências estéreo utilizadas nos testes, com resolução 1024×768 .

A sequência *Book Arrival* [36] apresenta um ambiente interior do tipo estúdio, com duas pessoas no plano principal em interação e movimento moderado. O conteúdo de

fundo é estático com um detalhe elevado. A câmara encontra-se em posição fixa. Esta sequência apresenta uma estrutura de profundidade com complexidade média, como se pode observar no exemplo de uma imagem desta sequência na Figura 5.3 (c).

A sequência *Ballroom*, obtida em ambiente interior, apresenta detalhe médio, movimento complexo de objetos, reflexões, estrutura de profundidade complexidade média e câmara fixa. As condições de captura de vídeo incluem câmaras alinhadas horizontalmente com a distância entre os eixos óticos paralelos de 19.5 cm e a uma altura de 1.5 m do plano inferior [37].

A sequência *Exit*, obtida em ambiente interior, apresenta detalhe médio, movimento complexo de objetos, reflexões, estrutura de profundidade complexa e câmara fixa. As condições de captura são as mesmas da sequência *Ballroom*.

As Figuras 5.4 (a) e (b) ilustram um exemplo de uma imagem para as sequências *Ballroom* e *Exit*, respetivamente.



(a) *Ballroom* (Câmara 4, Imagem 1).

(b) *Exit* (Câmara 4, Imagem 1).

Figura 5.4: Sequências estéreo utilizadas nos testes, com resolução 640×480 .

5.2 Determinação de parâmetros

Um parâmetro comum a ambos os métodos em comparação é a área útil da imagem auxiliar a melhorar. Devido à distância entre as câmaras, parte do conteúdo da imagem auxiliar junto a uma das margens laterais (direita ou esquerda) não existirá na imagem base. Do mesmo modo, a zona junto a uma das margens da imagem base não se encontra na imagem auxiliar.

A imagem da vista com o conteúdo mais à direita foi selecionada, de forma arbitrária, como a vista base e a imagem com o conteúdo semelhante mais à esquerda como a vista auxiliar. Nos exemplos da Figura 5.5, as imagens em (a) e (d) correspondem à vista base e em (b) e (c) à vista auxiliar.

Deste modo, não é efetuado o processamento da faixa mais à direita da vista auxiliar já que esta não existe na vista base. Esta margem tem o tamanho em altura igual à altura da imagem e a mesma largura da janela de pesquisa utilizada na estimação de disparidade nos métodos desenvolvidos. Terá assim uma largura igual ou superior à disparidade existente entre cada vista, caso a largura da janela de pesquisa seja selecionada corretamente. Nas Figuras 5.5 (b) e (c) são ilustrados dois exemplos de vistas auxiliares onde são assinaladas as respectivas margens laterais direitas, neste caso de 48 *pixels* cada. Ignorando a margem direita da vista auxiliar, o restante conteúdo pode ser utilizado, em conjunto com a vista base, para estimar os detalhes a adicionar.

O tamanho do bloco quadrado centrado no vértice (*kernel*) utilizado para realizar a estimação de disparidade de cada vértice do triângulo é muito importante, já que irá influenciar a exatidão da operação. O mesmo se aplica à determinação do tamanho de bloco no caso do método baseado na pesquisa blocos, não para localização de vértices, mas para localização do conteúdo de toda a imagem particionada em blocos quadrados iguais. Outro parâmetro muito importante em ambos os métodos é o tamanho da janela de pesquisa.

Como as sequências de vídeo utilizadas foram obtidas a partir de câmaras com uma disposição alinhada na horizontal, a altura da janela de pesquisa foi determinada pelo tamanho do bloco, adicionado de 4 *pixels* de tolerância, 2 *pixels* no topo e 2 *pixels* em baixo.

Uma possibilidade para determinar o melhor tamanho de bloco e largura da janela de pesquisa seria a realização de testes com a variação sequencial exaustiva de cada parâmetro. No entanto, não foi tomada essa opção por se tratar de um processo demasiado lento e provavelmente ter-se-ia chegado ao mesmo resultado. Assim, para determinar qual o tamanho do bloco de pesquisa são testados vários valores, escolhendo inicialmente uma janela de pesquisa suficientemente larga de modo a incluir a zona com o detalhe pretendido

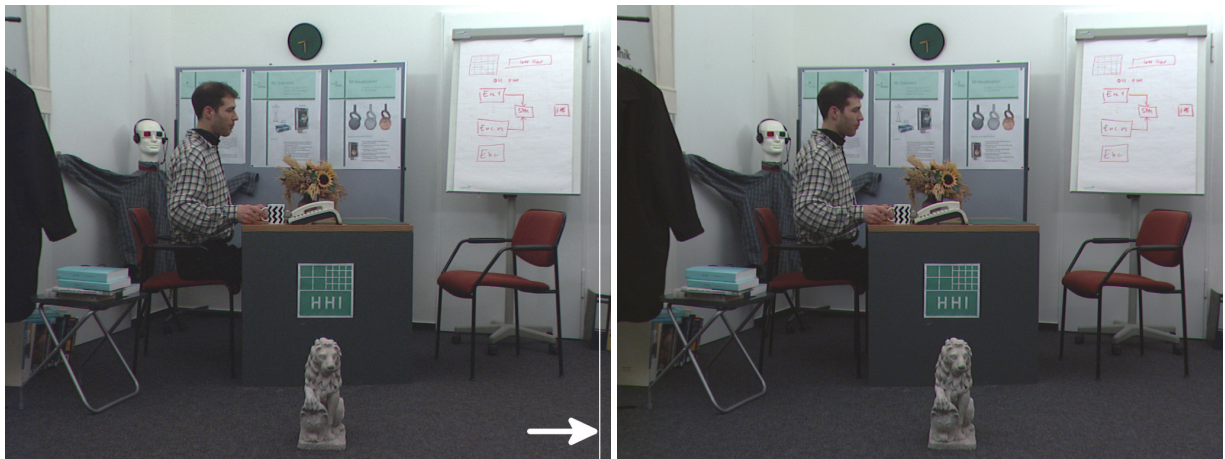
(a) *Balloons* (Câmara 4, Imagem 1).(b) *Balloons* (Câmara 5, Imagem 1).(c) *Book Arrival* (Câmara 10, Imagem 1).(d) *Book Arrival* (Câmara 11, Imagem 1).

Figura 5.5: Exemplos de identificação da margem da vista auxiliar não existente na vista base.

na vista base. A escolha inicial da largura da janela tem também em conta a resolução da imagem e disparidade observada. Quanto maior a resolução, maior a janela de pesquisa utilizada.

Após várias iterações de testes com variação do tamanho do bloco e janela de pesquisa, a largura da janela de pesquisa que produz resultados mais favoráveis é utilizada nos testes finais para obter o tamanho do bloco. Por sua vez, com a obtenção do tamanho do bloco, este valor é utilizado para efetuar um último teste de variação da janela de pesquisa.

Numa primeira fase os testes foram realizados de modo a determinar o melhor tamanho de bloco, *kernel* e janela de pesquisa para todas as sequências utilizadas [38]. No entanto, com a realização de testes adicionais com um maior número de sequências estéreo de várias resoluções, observa-se que o melhor tamanho de bloco e janela de pesquisa tende a ser semelhante para imagens com a mesma resolução e por vezes diferente para outras resoluções. Os parâmetros obtidos não podem ser considerados comprovadamente os

mesmos para todas as sequências com a mesma resolução, devido ao reduzido número de sequências diferentes utilizadas com a mesma resolução. No entanto, procurou-se selecionar um tamanho de bloco para cada conjunto de sequências com a mesma resolução, já que a seleção de um único valor para o tamanho de bloco e janela de pesquisa iria beneficiar apenas algumas sequências. Estes testes iniciais são realizados utilizando as primeiras cinco imagens de cada sequência em teste.

Para cada conjunto de sequências de cada resolução, após a atribuição da respectiva janela de pesquisa, são efetuados os testes de variação do tamanho de bloco. No caso do tamanho do *kernel* no método baseado na malha triangular, é utilizado um tamanho que varia entre 3 e 49 *pixels* para as sequências de maior resolução e entre 3 e 37 *pixels*, para as restantes. A variação do *kernel* é realizada utilizando apenas tamanhos ímpares, de modo que o vértice permaneça sempre no centro do bloco.

No caso do método baseado em pesquisa de blocos, a variação do tamanho de bloco quadrado situa-se entre 2 e 50 *pixels* para sequências de elevada resolução e entre 2 e 40 *pixels* para as restantes. Neste caso, a variação utiliza apenas os números pares do intervalo, de forma a dividir a totalidade da imagem em blocos com o mesmo tamanho.

Nos métodos em teste, para o tamanho de *kernel* e bloco foi selecionado inicialmente um intervalo de variação menor que o descrito. No entanto, devido à obtenção de resultados muito próximos, o intervalo foi sendo progressivamente aumentado de forma a visualizar claramente a evolução dos resultados com essa variação, chegando assim aos valores atuais descritos. A Tabela 5.2 mostra os intervalos de teste atribuídos a cada conjunto de imagens agrupadas por resolução.

Tabela 5.2: Intervalo de teste do tamanho do *kernel* e bloco (*pixels*).

Resolução da sequência	Mét. da Malha	Mét. Pesquisa Bloco
1920 × 1088	3 × 3 a 49 × 49	2 × 2 a 50 × 50
1280 × 960	3 × 3 a 37 × 37	2 × 2 a 40 × 40
1024 × 768	3 × 3 a 37 × 37	2 × 2 a 40 × 40
640 × 480	3 × 3 a 37 × 37	2 × 2 a 40 × 40

A análise dos resultados dos vários testes obtidos para um conjunto de sequências com uma determinada resolução exemplifica o estudo efetuado com os restantes conjuntos de sequências com outras resoluções. São disponibilizados no Apêndice A os resultados obtidos com a variação de tamanho de *kernel* (Secção A.1), bloco (Secção A.3) e janela de pesquisa (Secção A.2 e A.4), para cada sequência de teste nas várias subamostragens testadas com a vista auxiliar.

Tomando como exemplo os resultados obtidos com as sequências de tamanho 1920 ×

1088 *pixels*, a aplicação da variação do tamanho do *kernel* no método baseado na malha triangular, evidencia que um *kernel* entre 23 e 43 *pixels* será a melhor solução para estas sequências. Deste modo, foi selecionado um valor intermédio para o tamanho do *kernel* de 33 *pixels*, tendo em conta, no entanto, que a seleção de um valor menor iria beneficiar a diminuição complexidade computacional.

Apesar de possuírem a mesma resolução, o conteúdo das duas sequências em teste é muito diferente, tal como descrito na Secção 5.1. O elevado detalhe na sequência *GT-Fly* permite que o processamento, tanto a nível de disparidade dos vértices da malha triangular como da quantidade de detalhes que podem ser adicionados, mostre uma evolução clara com a alteração do tamanho de *kernel* utilizado. Na Figura 5.6 podem-se observar os resultados obtidos para cada QP, com uma janela de pesquisa de 100 *pixels* e vista auxiliar subamostrada na horizontal e vertical simultaneamente. Para cada *kernel* testado observa-se a diferença, em dB, entre a qualidade objetiva (PSNR – *Peak Signal-to-Noise Ratio*) da imagem processada com o método e o PSNR da imagem não processada.

Neste caso, a utilização de um tamanho de bloco relativamente grande, comparativamente ao *kernel* obtido para sequências de menor resolução, pode ser justificado pela maior quantidade de informação necessária para que o conteúdo do bloco no vértice seja mais facilmente distinguível do respetivo conteúdo onde está inserido.

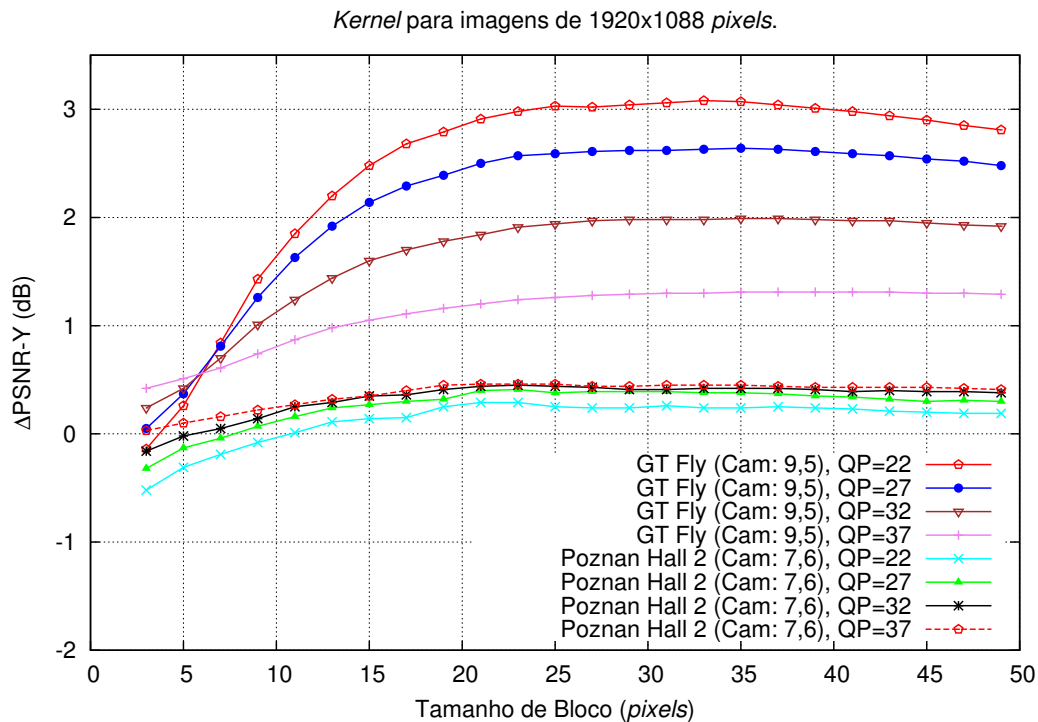


Figura 5.6: Diferença de PSNR da vista auxiliar, das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do *kernel*, com a vista auxiliar subamostrada na horizontal e vertical.

Após a seleção do tamanho de bloco / *kernel* para cada conjunto de sequências, segue-se o teste da largura da janela de pesquisa. No caso do método baseado na malha triangular, para as sequências de 1920×1088 *pixels*, foram efetuados testes com largura entre 38 e 208 *pixels*. O teste com sequências de 1280×960 e 640×480 *pixels*, incluiu valores entre 28 e 148 *pixels*. As restantes sequências de 1024×768 *pixels*, utilizaram valores entre 28 e 88 *pixels* para a largura da janela.

No método baseado em pesquisa blocos são utilizados os intervalos 40 a 210 *pixels* para o teste da largura da janela de pesquisa em sequências de 1920×1088 *pixels*. Para as restantes sequências são testados os valores no intervalo de 30 a 150 *pixels*.

Tal como no caso do teste de tamanho de *kernel* e bloco, os testes para encontrar o melhor valor para a largura da janela de pesquisa em cada método, foram realizados aumentando progressivamente o tamanho da janela. Neste caso a variação é realizada de 10 em 10 *pixels* de modo a verificar de forma clara a evolução dos resultados com a aplicação de janelas com tamanhos diferentes, obtendo-se assim os intervalos de teste apresentados na Tabela 5.3.

Tabela 5.3: Intervalo de teste da largura da janela de pesquisa (*pixels*).

Resolução da sequência	Mét. da Malha	Mét. Pesquisa Bloco
1920×1088	38 a 208	40 a 210
1280×960	28 a 148	30 a 150
1024×768	28 e 88	30 a 150
640×480	28 a 148	30 a 150

Retomando o exemplo das sequências com 1920×1088 *pixels*, as sequências *GT Fly* e *Poznan Hall 2* apresentam resultados mais favoráveis para janelas de pesquisa com uma largura superior a 78 *pixels*. Na Figura 5.7 observa-se as curvas obtidas para a variação da largura da janela de pesquisa para cada QP. Com sequências de resolução elevada o aumento da largura da janela não tem praticamente outro efeito a partir do valor verificado. Esta situação deve-se à combinação do tamanho do *kernel* utilizado com a elevada resolução da imagem que permite distinguir com mais facilidade o vértice pretendido, mesmo numa área mais alargada, relativamente a imagens de baixa resolução. Tendo em conta a análise efetuada foi escolhido uma largura de 100 *pixels* para este conjunto de sequências.

O processo para obtenção dos parâmetros para os restantes conjuntos de sequências é semelhante ao exemplificado. As curvas resultantes podem ser consultadas no Apêndice A, para o caso do método baseado na malha triangular (Secção A.2) e método baseado na pesquisa de blocos (Secção A.4).

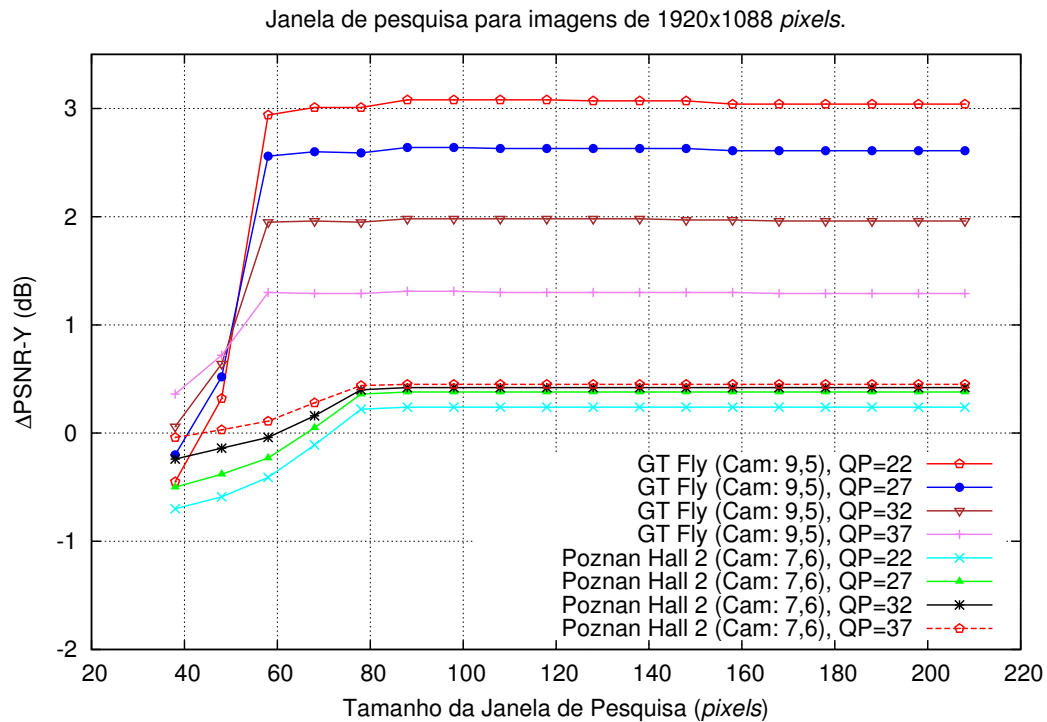


Figura 5.7: Diferença de PSNR da vista auxiliar, das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa no método da malha triangular, com a vista auxiliar subamostrada na horizontal e vertical.

Com a determinação dos valores de tamanho de *kernel*, bloco e largura das janelas de pesquisa para os vários conjuntos de sequências, discriminados na Tabela 5.4, pode-se prosseguir com a comparação entre o método baseado na malha triangular e pesquisa de blocos. No entanto, antes desse passo é ainda necessário efetuar a comparação entre a utilização da vista auxiliar ou base como vista de referência para procurar os blocos e vértices correspondentes da vista base ou auxiliar, respetivamente. Esta comparação é realizada na Secção 5.3.

Tabela 5.4: Resultados obtidos para o tamanho de bloco e janela de pesquisa (*pixels*).

Resolução da sequência	Método da Malha		Método Pesquisa Bloco	
	<i>Kernel</i>	Janela	Bloco	Janela
1920 × 1088	33 × 33	100 × 37	32 × 32	90 × 36
1280 × 960	15 × 15	48 × 19	8 × 8	50 × 12
1024 × 768	13 × 13	148 × 17	14 × 14	50 × 18
640 × 480	9 × 9	48 × 13	2 × 2	50 × 6

5.3 Imagem de referência na pesquisa de bloco

Os testes para determinação do tamanho de bloco, no método baseado na malha triangular e método baseado na pesquisa de blocos, são realizados recorrendo à estimação de disparidade de cada bloco da vista auxiliar na respetiva janela de pesquisa da vista base. Deste modo, a vista auxiliar é a vista com os blocos de referência a procurar na outra vista, tal como descrito no Capítulo 4.

No entanto, também existe a possibilidade de obter a estimação de disparidade entre as imagens através da pesquisa de bloco ou *kernel* da vista base numa janela de pesquisa na vista auxiliar. Nesse caso a vista com os blocos de referência será a vista base. No caso do método da malha isso implica a construção da malha triangular na vista base de modo a obter os vértices para as coordenadas dos blocos a pesquisar na outra vista. No caso do método baseado na pesquisa de blocos, será a vista base a ser particionada em blocos quadrados iguais para pesquisar na vista auxiliar.

Uma diferença que poderá surgir com a utilização da vista base como vista de referência residirá na maior propensão para obtenção de erros de pesquisa, devido à utilização da janela de pesquisa na vista auxiliar em vez da vista base. Devido à remoção dos detalhes da vista auxiliar resultante da subamostragem, cada imagem desta sequência irá apresentar valores em cada bloco mais homogêneos do que os correspondentes blocos na vista base.

A escolha da vista de referência poderá assim afetar a precisão da operação dos métodos de processamento. De modo a verificar o impacto que a seleção da vista de referência exerce no resultado final do processamento, neste caso do método baseado na malha triangular, são realizados testes tendo a vista base como referência.

Com a substituição da vista de referência, de modo a tornar a comparação coerente com os testes já realizados, torna-se necessário a determinação dos parâmetros de tamanho de bloco para esta nova configuração. Deste modo são efetuados testes semelhantes aos efetuados na Secção 5.2 substituindo a vista de referência.

Recorrendo ao exemplo dos testes com a sequência *Champagne Tower*, neste caso para a vista auxiliar subamostrada, simultaneamente na horizontal e vertical, pode-se comparar os resultados já obtidos anteriormente com os obtidos alterando a vista de referência. A comparação será efetuada aqui apenas como um caso exemplificativo da análise efetuada, sendo disponibilizado no Apêndice A, Secção A.5 as curvas obtidas com os resultados dos restantes testes realizados.

Os resultados obtidos para a sequência *Champagne Tower*, representados na Figura 5.8, permitem confirmar que a diferença de PSNR entre a vista auxiliar processada e a vista auxiliar não processada, em função do tamanho do *kernel*, para os dois casos da vista de referência, são semelhantes ou favoráveis à utilização da vista auxiliar como vista de

referência. De um modo geral, verificando todos os resultados obtidos pode-se confirmar que a utilização da vista base com a janela de pesquisa para encontrar cada *kernel* de referência da vista auxiliar dá melhores resultados do que a situação inversa.

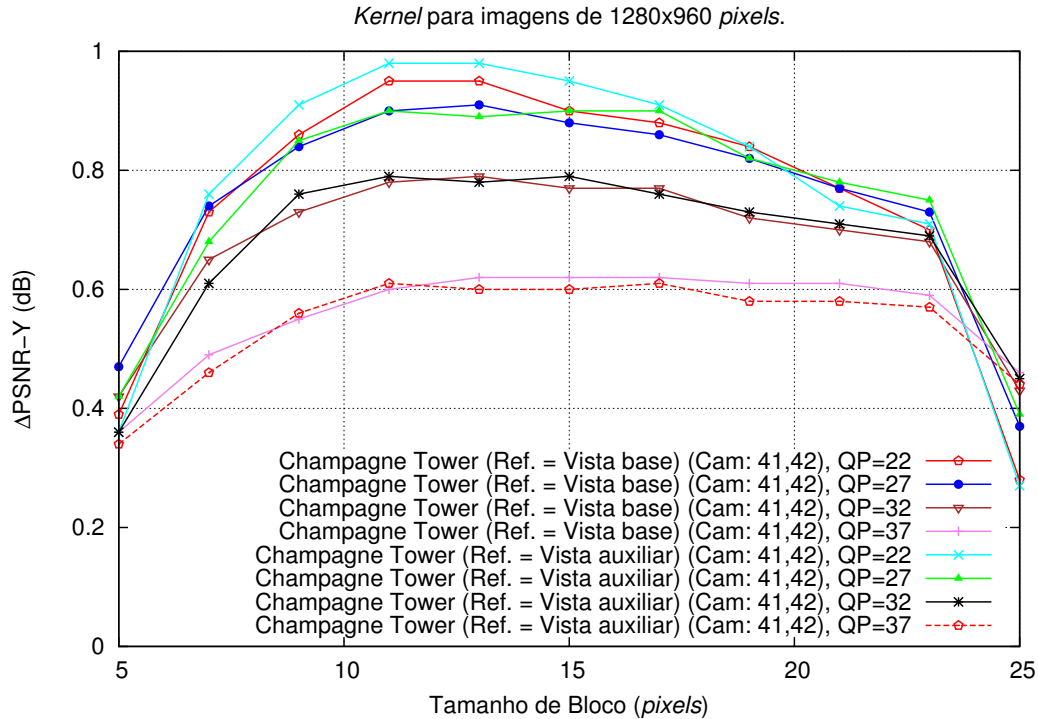


Figura 5.8: Diferença de PSNR da sequência *Champagne Tower* em função da variação do tamanho do *kernel*, comparando a utilização da vista auxiliar e base como vista de referência na pesquisa do *kernel*.

Na globalidade dos testes efetuados, nem sempre é possível obter resultados claros no que diz respeito à seleção da vista de referência. Isto acontece sobretudo em sequências com conteúdo com menor nitidez, isto é, menos detalhes, que resulta numa aproximação entre os resultados obtidos utilizando uma ou outra vista como referência. Nestas situações o conteúdo da janela de pesquisa será mais semelhante utilizando uma ou outra vista como vista de referência. Por outro lado, em sequências com detalhe muito elevado, como por exemplo a sequência *GT Fly*, apesar da subamostragem da vista auxiliar, o detalhe existente permite uma melhor exatidão na estimação de disparidade do *kernel* resultando num melhor desempenho se a vista base é a vista de referência.

De uma forma geral pode-se concluir que a utilização da vista auxiliar como vista de referência traz mais benefícios do que a utilização da vista base. Assim, tendo em conta os resultados obtidos para o teste do tamanho do *kernel* no método baseado na malha triangular, utilizando uma ou outra vista como vista de referência, não se verifica a necessidade de testes adicionais. Para os restantes testes a efetuar será definida como a vista de referência a vista auxiliar, enquanto que a vista base terá a janela de pesquisa.

5.4 Adição da imagem de detalhes

O processamento do método baseado na malha triangular termina com a adaptação e adição dos detalhes do conteúdo da imagem da vista base ao conteúdo da imagem da vista auxiliar. No caso do método baseado na pesquisa de blocos, o processamento termina com a adição dos detalhes à vista auxiliar, de acordo com o resultado obtido na pesquisa efetuada. No Capítulo 4 é apresentada a técnica proposta para adição dos detalhes à imagem auxiliar, com base numa máscara construída a partir dos contornos da imagem. Pretendem-se comparar os resultados obtidos deste método de adição com a adição direta de detalhes.

A adição direta significa que após a adaptação dos detalhes do conteúdo da imagem da vista base ao conteúdo da imagem da vista auxiliar, os detalhes são adicionados bloco a bloco desde o início da imagem até ao final. No caso de utilização de uma máscara baseada nos contornos da imagem, apenas os detalhes junto aos contornos da imagem são adicionados à imagem auxiliar. Justifica-se a utilização desta máscara desde que a imagem não apresente detalhes muito acentuados, pois nessa situação, a adição seletiva com base apenas nos contornos poderá não ser tão favorável como no caso da adição completa.

Para a realização dos testes das técnicas de adição de detalhes são utilizados os parâmetros obtidos anteriormente para a configuração do tamanho de bloco ou *kernel*. De uma forma geral o método baseado na malha triangular apresenta melhores resultados com a adição baseada na máscara de contornos, relativamente ao método baseado na pesquisa de blocos.

Tomando como exemplo os resultados obtidos para as sequências *Poznan Hall 2* e *GT Fly*, nas Figuras 5.9 e 5.10 observam-se as curvas com a relação entre o PSNR e a taxa de transmissão das sequências base codificada com o HEVC e auxiliar codificada com e sem subamostragem e pós processamento com adição direta ou baseada na máscara de contornos. A sequência *Poznan Hall 2* subamostrada na horizontal é um exemplo onde a adição dos detalhes com base na máscara de contornos favorece o resultado final. Já a sequência *GT Fly*, com a vista subamostrada na horizontal e vertical, exemplifica um caso onde a adição direta fornece melhor resultado, mas próximo do método proposto, sendo o elevado nível de detalhe uma característica desta sequência que deverá contribuir para uma boa localização de todos os vértices na imagem.

Com a execução de todos os testes com a vista auxiliar subamostrada na horizontal, vertical e ambas as direções, pode-se observar os resultados obtidos através dos gráficos disponibilizados no Apêndice A, Secção A.6. No entanto, a utilização da métrica *Delta Bjøntegaard* PSNR (BD-PSNR) e percentagem de *Delta Bjøntegaard* de taxa de trans-

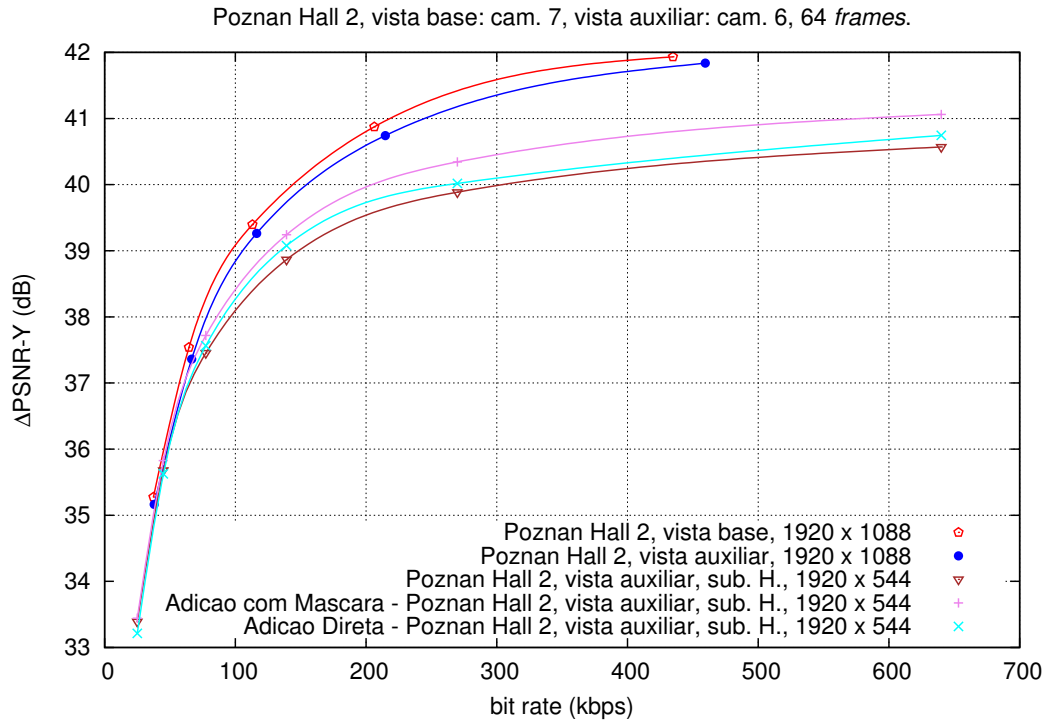


Figura 5.9: Comparação da taxa-distorção (RD) entre a adição direta e com base na máscara para a sequência *Poznan Hall 2* com a vista auxiliar subamostrada na horizontal.

missão ($\%BR - Bj\o ntgaard Delta Bit Rate$) [39] permite a análise dos resultados de uma forma compacta com a obtenção do ganho relativo entre os dois métodos em teste, em comparação com a não utilização do pós processamento. Quanto maior o BD-PSNR melhor o resultado da qualidade objetiva em relação à situação onde o método não é aplicado. Quanto menor (mais negativo) o $\%BR$ melhor será o resultado em relação à taxa de transmissão quando o método não é aplicado.

Na Tabela 5.5 pode-se observar um exemplo dos resultados obtidos com a métrica BD-PSNR para os dois modos de adição de detalhes, quando a vista auxiliar é subamostrada simultaneamente na horizontal e vertical e processada com o método baseado na malha triangular. Os restantes resultados, para as várias subamostragens com os dois métodos de melhoria da vista auxiliar, encontram-se no Apêndice A, Seção A.7

Os resultados obtidos evidenciam a capacidade que os métodos de melhoria da vista auxiliar têm para localizar corretamente os blocos desta vista na vista base. Quando os detalhes da imagem são acentuados, a pesquisa efetuada nos métodos de processamento da vista auxiliar terá necessariamente mais facilidade em encontrar o *kernel* dos vértices ou blocos semelhantes na outra vista. Assim, com uma boa localização dos blocos, a adição direta dos detalhes poderá compensar o eventual ruído gerado com essa operação. Enquanto que uma adição localizada apenas nos contornos irá ignorar blocos de detalhes corretamente localizados em zonas fora desses contornos.

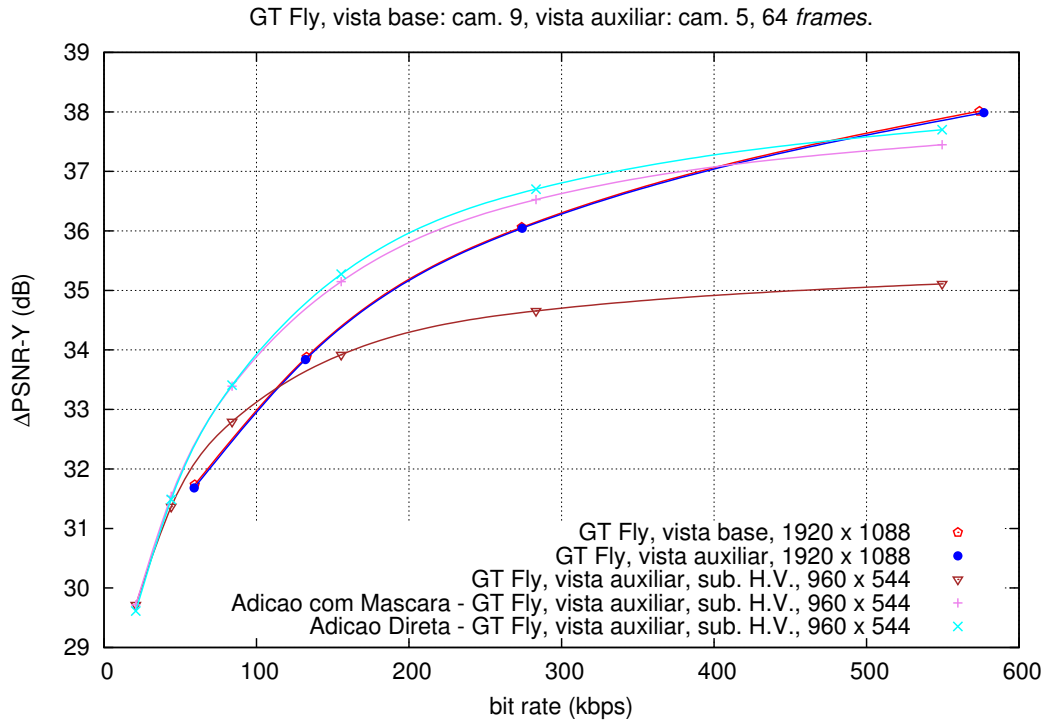


Figura 5.10: Comparação da taxa-distorção (RD) entre a adição direta e com base na máscara para a sequência *GT Fly* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.

Tabela 5.5: Avaliação do método de adição de detalhes do método da malha triangular, com vista auxiliar subamostrada na horizontal e vertical.

Sequência	BD-PSNR		%BR	
	Máscara	Direta	Máscara	Direta
<i>GT Fly</i>	1.547	1.695	-45.740	-47.421
<i>Poznan Hall 2</i>	0.412	0.335	-18.604	-14.516
<i>Pantomime</i>	0.616	0.506	-13.546	-11.536
<i>Champagne Tower</i>	0.605	0.593	-16.134	-15.873
<i>Balloons</i>	0.410	0.380	-9.409	-8.739
<i>Kendo</i>	0.371	-0.774	-8.235	22.677
<i>Book Arrival</i>	1.112	1.098	-28.512	-28.537
<i>Ballroom</i>	0.048	-0.007	-1.144	0.331
<i>Exit</i>	-0.148	-0.255	5.755	9.875

Por outro lado, seguindo o mesmo raciocínio com sequências com um nível de detalhe menos acentuado, a adição completa de todos os detalhes não deverá ser tão favorável

como no caso de utilização de máscara baseada nos contornos da imagem. Neste caso o ruído gerado pela operação de adição em zonas com conteúdo mais homogêneo poderá justificar a utilização da máscara para adição de detalhes apenas nos contornos.

Globalmente, o método baseado na malha triangular apresenta resultados mais favoráveis com a adição baseada na máscara de contornos, relativamente ao método baseado na pesquisa de blocos. A máscara limita a adição de detalhes aos blocos localizados junto dos contornos. Estas são as zonas da imagem onde o método baseado na malha triangular desenvolve os algoritmos de localização de contornos e cantos, triangulação e pesquisa de cada *kernel* com os vértices. A localização dos vértices em zonas com elevado detalhe permite uma melhoria da imagem auxiliar principalmente junto aos contornos. O método baseado na pesquisa de blocos funcionará igualmente bem nas zonas de elevado detalhe, no entanto, poderá não localizar corretamente toda a zona envolvente aos contornos, mas localizar corretamente outros detalhes não identificados como contornos. Os resultados obtidos justificam a utilização deste método de adição de detalhes com recurso à máscara baseada nos contornos da imagem.

5.5 Discussão de resultados

Com a seleção dos parâmetros de configuração e seleção da técnica de adição de detalhes da máscara baseada nos contornos da imagem, resta utilizar os dados já obtidos para comparar os resultados entre o método baseado na malha triangular e método baseado na pesquisa de blocos, para todas as sequências em teste. De modo a determinar se a utilização dos métodos desenvolvidos superam a não utilização da codificação assimétrica torna-se necessário comparar ambos os resultados obtidos.

Com a técnica de adição de detalhes com base na máscara, a utilização do método baseado na malha triangular apresenta um resultado mais favorável do que o método baseado na pesquisa de blocos. Na Figura 5.11 pode-se observar as curvas com a relação entre o PSNR e a taxa de transmissão das sequências base codificada com o HEVC e auxiliar codificada com e sem subamostragem e aplicação de cada método de melhoria da vista auxiliar. O elevado nível de detalhe que esta sequência apresenta facilita a correta localização dos vértices da imagem, permitindo assim que a adaptação dos detalhes da vista base à auxiliar seja efetuada sem gerar demasiado ruído.

Uma das sequências testadas apresenta uma situação muito particular, onde a utilização do método da pesquisa de bloco não exerce qualquer benefício no resultado final, independentemente do tipo de subamostragem da vista auxiliar. Para exemplificar, na Figura 5.12 pode-se observar as curvas obtidas para a sequência *Pantomime* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical. Devido ao tipo de con-

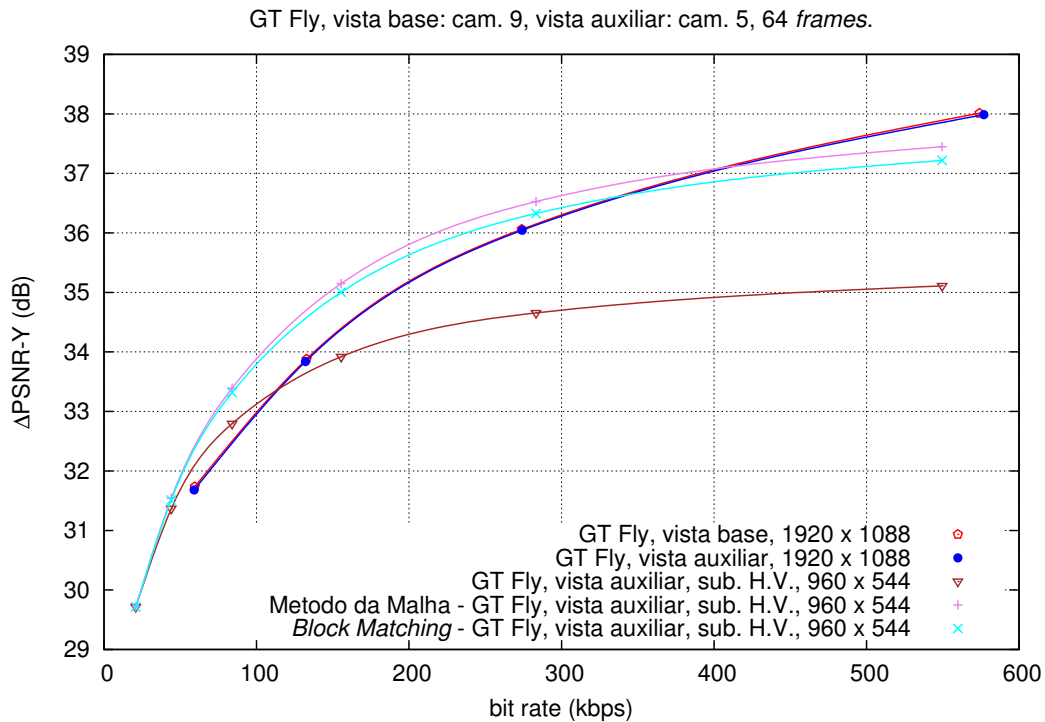


Figura 5.11: Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência *GT Fly* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.

teúdo desta sequência, todos os blocos que melhoram a imagem são anulados pelo ruído produzido pelos restantes blocos localizados incorretamente. O método baseado na malha triangular produz aqui uma melhoria significativa na vista auxiliar em comparação com a pesquisa de blocos. Esta situação deve-se ao nível de detalhe da sequência e à utilização de vértices colocados em pontos estratégicos, fáceis de localizar.

O método baseado na malha triangular por vezes apresenta resultados semelhantes aos obtidos com o método baseado na pesquisa de blocos. É exemplo disso a sequência *Poznan Hall 2* que, apesar da elevada resolução, não apresenta o mesmo nível de nitidez da sequência *GT Fly*. Na Figura 5.13 observam-se as curvas RD obtidas para a vista auxiliar subamostrada simultaneamente na horizontal e vertical. O nível de detalhe que esta sequência apresenta não é o suficiente para que o método baseado na malha triangular consiga obter uma boa localização dos vértices da malha, de modo a adaptar corretamente uma maior quantidade de detalhes da vista base à vista auxiliar.

Outro aspeto a considerar é a dependência da resolução espacial. Para sequências com baixa resolução espacial, os ganhos de qualidade na melhoria da imagem são reduzidos, já que o conteúdo da imagem fornece menos detalhes de alta frequência, logo os detalhes obtidos da imagem de resolução original tem menos informação. Esta situação dificulta também a correta localização dos vértices no caso do método baseado na malha triangular.

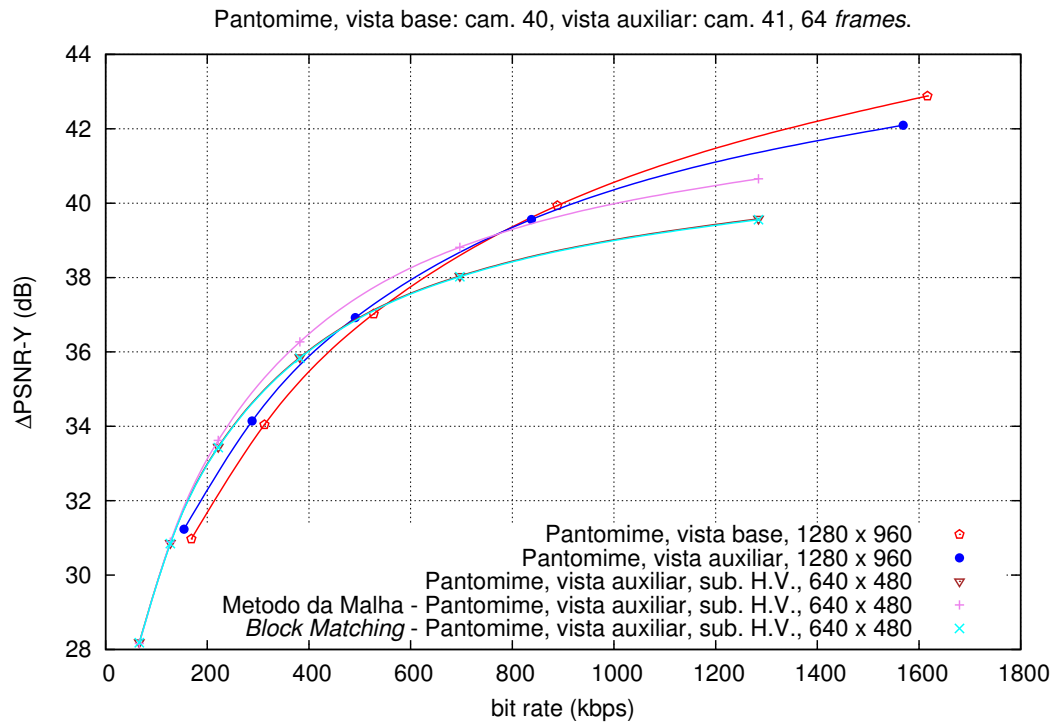


Figura 5.12: Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência *Pantomime* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.

Na Seção A.8 do Apêndice A são disponibilizados os restantes resultados de todas as sequências, obtidos sob a forma de gráfico com a relação entre o PSNR e a taxa de transmissão, para a vista auxiliar processada com os métodos considerados. De modo a resumir e a comparar mais facilmente os resultados obtidos nos vários testes, as Tabelas 5.6 a 5.8, indicam o ganho relativo BD-PSNR e percentagem de ganho de taxa de transmissão entre os dois métodos propostos, em comparação com o não processamento da imagem (sobreamostragem simples).

Ambos os métodos permitem aumentar a qualidade objetiva da imagem auxiliar na maior parte das sequências testadas. No entanto, o método baseado na malha triangular obtém, quase sempre, melhores resultados em relação ao método baseado na pesquisa de blocos. Os resultados obtidos permitem assim comprovar que um método que depende exclusivamente da pesquisa de blocos para localizar os detalhes a adicionar à imagem auxiliar, terá tendência para anular o efeito da melhoria realizada com o ruído gerado, devido à incorreta localização de uma parte dos blocos na imagem. O método da malha triangular, sendo um método que depende principalmente da correta localização de alguns pontos situados em zonas de elevado detalhe, apresenta resultados mais favoráveis que um método baseado apenas em pesquisa de blocos.

Os testes efetuados na vista auxiliar usaram subamostragem nas direções horizontal

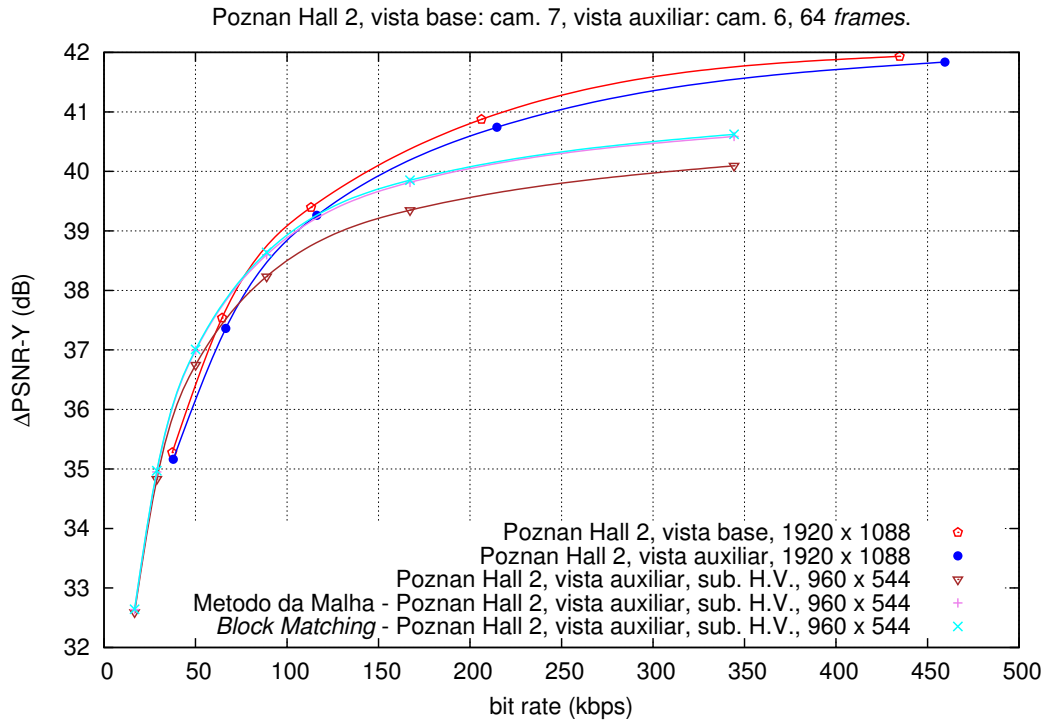


Figura 5.13: Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência *Poznan Hall 2* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.

Tabela 5.6: Avaliação dos métodos de melhoria da vista auxiliar subamostrada na horizontal (H).

Sequência	BD-PSNR		%BR	
	Malha	Bloco	Malha	Bloco
<i>GT Fly</i>	0.912	0.769	-36.087	-33.053
<i>Poznan Hall 2</i>	0.419	0.480	-21.447	-23.658
<i>Pantomime</i>	0.709	-0.021	-16.382	0.677
<i>Champagne Tower</i>	0.566	0.390	-16.066	-11.897
<i>Balloons</i>	0.349	0.278	-8.949	-7.342
<i>Kendo</i>	0.307	0.269	-7.338	-6.494
<i>Book Arrival</i>	0.602	0.627	-17.708	-18.319
<i>Ballroom</i>	-0.016	-0.011	0.463	0.279
<i>Exit</i>	-0.163	-0.008	7.002	0.351

(H) e vertical (V) independentemente, assim como em ambas as direções (H/V) simultaneamente. Os resultados mostram que com a subamostragem H/V são alcançados ganhos

Tabela 5.7: Avaliação dos métodos de melhoria da vista auxiliar subamostrada na vertical (V).

Sequência	BD-PSNR		%BR	
	Malha	Bloco	Malha	Bloco
<i>GT Fly</i>	1.560	1.503	-46.285	-44.508
<i>Poznan Hall 2</i>	0.043	0.006	-2.151	-0.130
<i>Pantomime</i>	0.065	-0.005	-1.559	0.117
<i>Champagne Tower</i>	0.242	0.055	-7.280	-1.698
<i>Balloons</i>	0.203	0.119	-4.890	-2.953
<i>Kendo</i>	0.199	0.138	-4.917	-3.446
<i>Book Arrival</i>	0.987	0.956	-26.426	-25.809
<i>Ballroom</i>	0.071	-0.012	-1.814	0.319
<i>Exit</i>	-0.064	-0.004	2.804	0.167

Tabela 5.8: Avaliação dos métodos de melhoria da vista auxiliar subamostrada na horizontal e vertical (H/V).

Sequência	BD-PSNR		%BR	
	Malha	Bloco	Malha	Bloco
<i>GT Fly</i>	1.547	1.378	-45.740	-43.038
<i>Poznan Hall 2</i>	0.412	0.447	-18.604	-19.955
<i>Pantomime</i>	0.616	-0.021	-13.546	0.602
<i>Champagne Tower</i>	0.605	0.364	-16.134	-10.453
<i>Balloons</i>	0.410	0.318	-9.409	-7.576
<i>Kendo</i>	0.371	0.313	-8.235	-7.021
<i>Book Arrival</i>	1.112	1.100	-28.512	-28.438
<i>Ballroom</i>	0.048	-0.014	-1.144	0.380
<i>Exit</i>	-0.148	-0.008	5.755	0.298

superiores relativamente à utilização das outras subamostragens (H e V) em separado. Tendo a vista auxiliar nesta situação uma menor quantidade de informação devido à subamostragem em ambas as direções horizontal e vertical, a informação com os detalhes disponível na vista base é superior, pois estes são gerados a partir da subtração da vista base com a sua versão subamostrada.

Com a vista auxiliar subamostrada simultaneamente na horizontal e vertical, para as seqüências *Balloons*, *Kendo*, *Pantomime* e especialmente *GT Fly*, é obtida uma taxa de transmissão abaixo da qual vale a pena utilizar codificação assimétrica de vídeo estéreo, já que são atingidas menores taxas de transmissão na agregação de ambas as vistas (*simulcast*), assumindo que a vista auxiliar pode ser codificada com menor qualidade. Esta situação é ilustrada, respetivamente, nas Figuras 5.14, 5.15, 5.12 e 5.11.

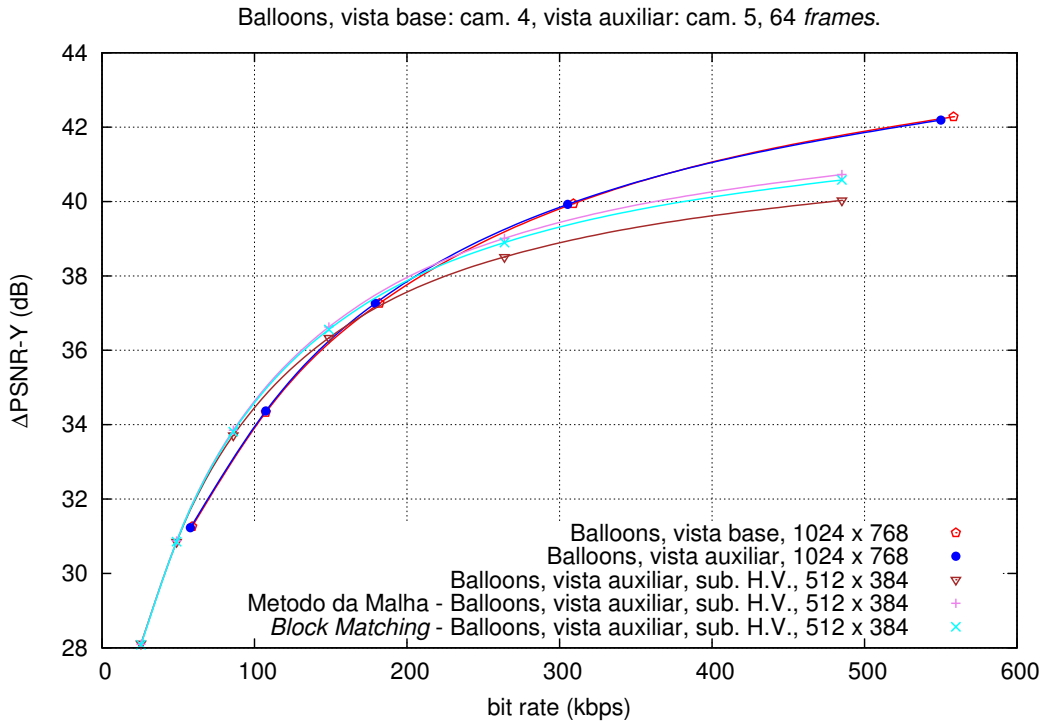


Figura 5.14: Comparação da taxa-distorção (RD) utilizando os métodos propostos com a seqüência *Balloons* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.

A avaliação de qualidade subjetiva de ambos os métodos é aqui exemplificada através da análise e observação de duas amostras retiradas da seqüência *Book Arrival*. Nas Figuras 5.16 (a) a (l) é ilustrada a primeira amostra nas várias fases da análise, obtida a partir da vista da Câmara 10, Imagem 1, da seqüência *Book Arrival*.

A imagem original da amostra, ou seja, antes da subamostragem e codificação, está representada na Figura 5.16 (a). A figura em baixo, Figura 5.16 (e), representa o resultado obtido com o cálculo da diferença efetuado entre a vista auxiliar original e a mesma amostra decodificada sem prévia subamostragem, com um aumento dos *pixels* para metade do valor máximo da intensidade de pixel de modo a tornar visível os detalhes obtidos. Uma alternativa ao cálculo simples da diferença é a diferença logarítmica, que permite acentuar os detalhes obtidos. Neste caso a Figura 5.16 (i), representa essa diferença. Observando as imagens com a representação da diferença, verifica-se, como seria de esperar,

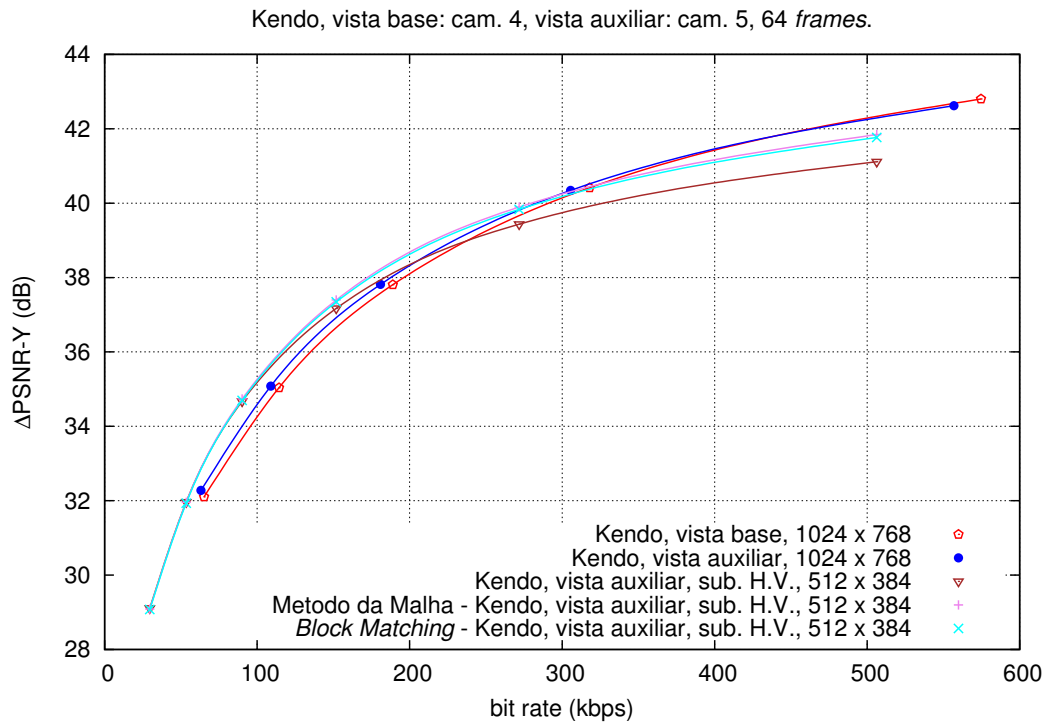


Figura 5.15: Comparação da taxa-distorção (RD) utilizando os métodos propostos com a sequência *Kendo* com a vista auxiliar subamostrada simultaneamente na horizontal e vertical.

que a diferença não resulta em detalhes visíveis porque a imagem a subtrair não sofreu o processo de subamostragem, isto é, mantém grande parte dos detalhes após a codificação

Subamostrando a imagem, seguido da respectiva codificação e decodificação e sobreamostragem a imagem irá perder uma parte significativa dos seus detalhes, resultando numa imagem desfocada. A Figura 5.16 (b) representa a amostra obtida, que resultou numa qualidade objetiva de 33.40 dB. Subtraindo a amostra nesta fase ao correspondente bloco de imagem na vista auxiliar decodificada mas sem prévia subamostragem, obtém-se a imagem representativa na Figura 5.16 (f) e (j) para versão da diferença logarítmica. Neste caso, com a perda notória de detalhes, estes são facilmente visíveis nos resultados obtidos com a diferença.

Aplicando o método baseado na malha triangular, será de esperar melhorias no detalhe da imagem. Isso mesmo pode ser observado na Figura 5.16 (c), onde a imagem do bloco da amostra ficou visivelmente mais nítida em relação ao não processamento observado no passo anterior, conseguindo uma qualidade objetiva de 36.17 dB. Tal como anteriormente são representados nas Figuras 5.16 (g) e (k) a diferença linear e logarítmica entre a imagem da vista auxiliar decodificada sem prévia subamostragem e a imagem após o processamento do método baseado na malha triangular. Observa-se ainda alguns detalhes que não foram adicionados com o método de melhoria. No entanto, comparativamente ao

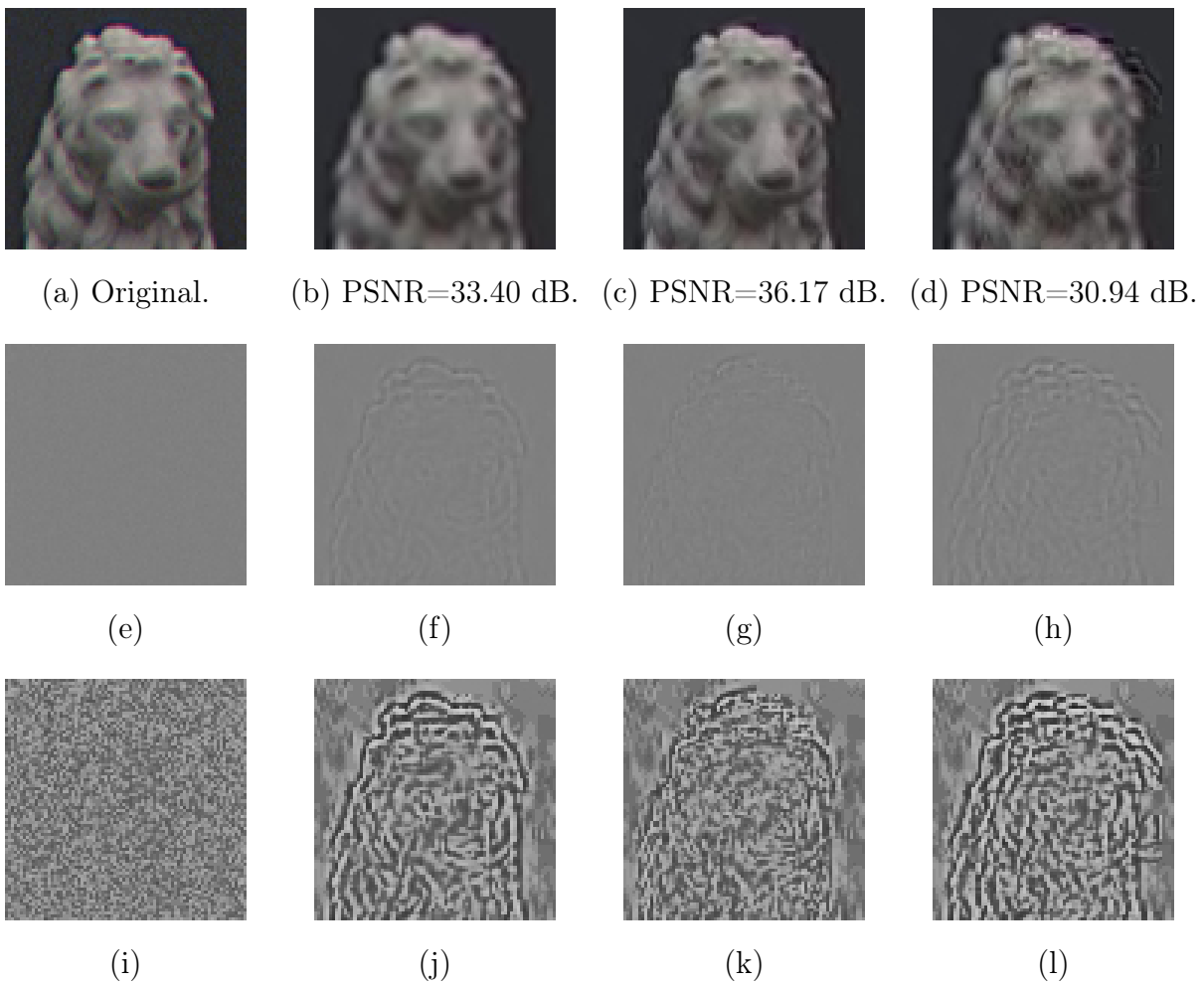


Figura 5.16: Comparação subjetiva e objetiva num detalhe da Imagem 1 de *Book Arrival* (Cam. 10, Imagem 1) e diferenças linear e logarítmica entre as amostras obtidas e a amostra de referência.

não processamento, este método reduz significativamente a quantidade de detalhes que ficaram por adicionar.

Utilizando o método baseado na pesquisa de blocos, na Figura 5.16 (d) observa-se que nesta amostra este processo resulta em detalhes deslocados, indicando um cálculo incorreto da disparidade nesta zona. Esta situação reduz a qualidade da imagem obtida, quer subjetivamente mesmo em relação ao não processamento de melhoria com o método, quer objetivamente, obtendo um PSNR de 30.94 dB. Como seria de esperar as diferenças obtidas mostram uma quantidade elevada de detalhes que não foram adicionados corretamente à imagem. As Figuras 5.16 (h) e (l) representam as diferenças linear e logarítmica, obtidas entre a imagem da vista auxiliar decodificada sem prévia subamostragem e a imagem após o processamento do método baseado na pesquisa de blocos.

A segunda amostra é analisada de forma semelhante à anterior. Neste caso foi selecionado outro bloco da imagem pertencente à mesma sequência, *Book Arrival*, obtida

a partir da vista da Câmara 10, Imagem 1. As várias fases da análise efetuada estão representadas nas Figuras 5.17 (a) a (l). Tratando-se de uma análise semelhante à efetuada anteriormente, serão aqui mencionados apenas os pormenores que diferenciam os resultados desta amostra da anterior.

A aplicação do método baseado na malha triangular resulta de uma significativa melhoria subjetiva da imagem, relativamente ao não processamento. No entanto, observando a Figura 5.17 (c), verifica-se que parte da imagem está distorcida. Observa-se que os detalhes na zona distorcida são colocados por cima de uma área que precisa receber esses detalhes, no entanto, a adaptação não foi efetuada corretamente. Esta situação pode ser justificada com o modo como a malha triangular é construída. O processo de triangulação de Delaunay, descrito na Secção 4.3 do Capítulo 4, evita o aparecimento de triângulos com ângulos internos muito pequenos. Neste caso, existe uma quantidade de detalhes no conteúdo da imagem, acumulados numa área relativamente pequena, dificultando a correta seleção de conteúdo com a malha triangular, para posterior adaptação com a transformada Afim.

Relativamente ao método de melhoria da imagem auxiliar baseado na pesquisa de blocos, observa-se melhorias em alguns pontos e distorção em parte da imagem da amostra, tal como se pode verificar na Figura 5.17 (d). Tratando-se de um conteúdo com secções semelhantes, a pesquisa de bloco pode falhar na correta identificação do local a aplicar os detalhes correspondentes. Neste caso ambos os métodos aplicam parte dos detalhes corretamente na imagem. Apesar do ruído gerado com a aplicação incorreta dos restantes detalhes, o resultado final da qualidade subjetiva e objetiva, é favorável à utilização de ambos os métodos, especialmente o método baseado na malha triangular.

Estes exemplos mostram dois casos distintos, onde é possível observar uma situação onde existe um benefício claro com a utilização do método proposto, assim como uma situação onde esse benefício não é tão relevante.

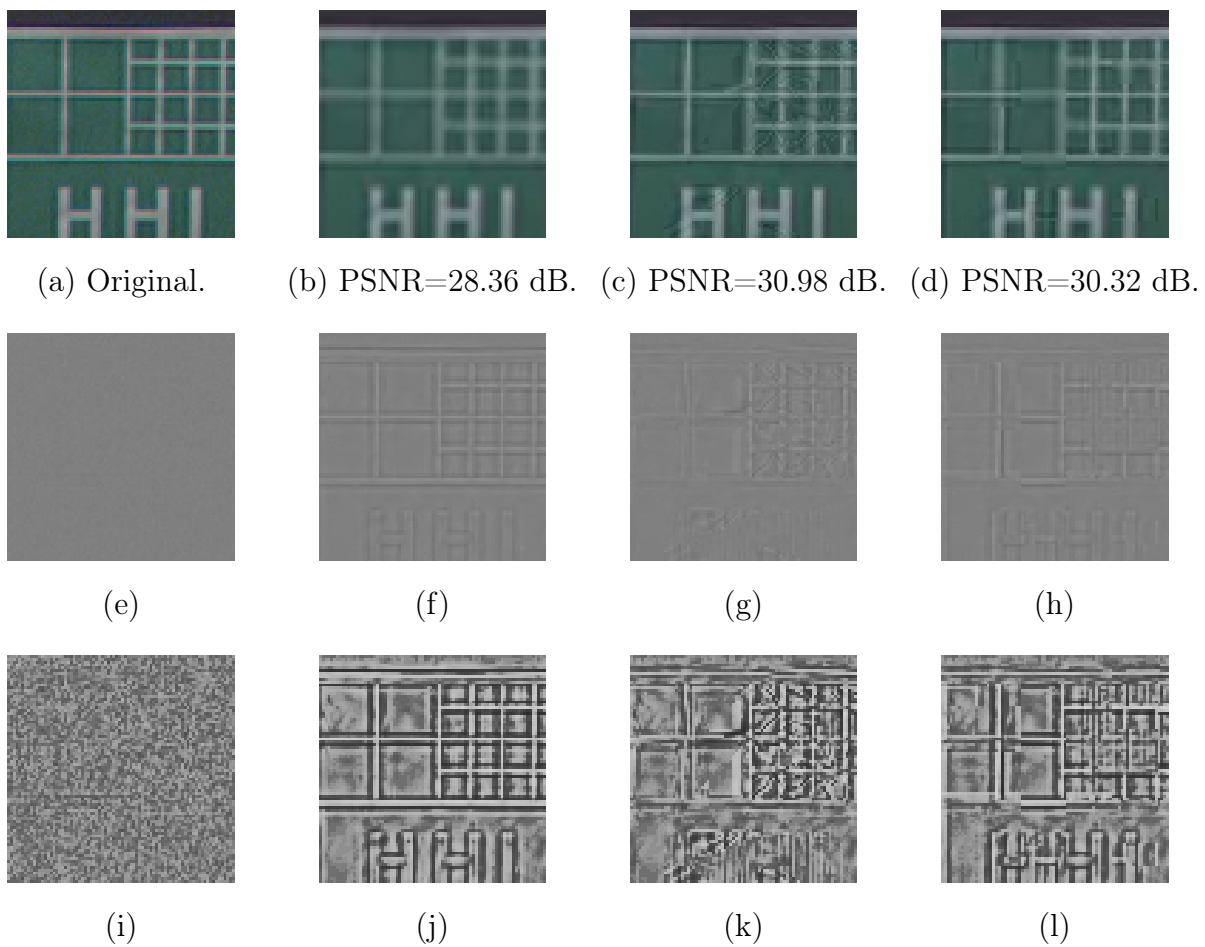


Figura 5.17: Comparação subjetiva e objetiva num detalhe da Imagem 1 de *Book Arrival* (Cam. 10, Imagem 1) e diferenças linear e logarítmica entre as amostras obtidas e a amostra de referência.

Capítulo 6

Conclusão e Trabalho Futuro

Nesta dissertação estudou-se a utilização de um método baseado numa malha triangular em conjunto com um método de adição de detalhes, para melhorar a qualidade da vista sobreamostrada numa transmissão estereoscópica de vídeo assimétrico. Os métodos propostos têm como objetivo a redução da taxa de transmissão da codificação de um par de vistas, necessárias para transmitir uma sequência estéreo.

Os testes realizados incluíram a comparação com um método baseado na pesquisa de blocos. Previamente aos testes de comparação entre os dois métodos, foram realizados testes para determinar o conjunto de parâmetros necessários que resultassem no melhor desempenho possível em ambos os métodos.

O tamanho de bloco utilizado para localizar os vértices revelou ser um parâmetro determinante no desempenho obtido com o método baseado na malha triangular. A utilização de um conjunto alargado de sequências com várias resoluções permitiu observar que os parâmetros obtidos, se determinados por subconjunto de sequências com a mesma resolução, resultariam num maior benefício para a melhoria da vista auxiliar, relativamente à utilização de um tamanho de bloco fixo para todas as sequências. No entanto, a baixa quantidade de sequências com a mesma resolução em teste não permite concluir que os tamanhos de bloco obtidos para os vários conjuntos de sequências, se apliquem a todas as sequências com a mesma resolução. Como trabalho futuro, será necessário realizar os mesmos testes sobre um conjunto alargado de sequências com a mesma resolução, de modo a determinar se existe um valor mais favorável à melhoria da vista auxiliar, dependente da resolução da imagem.

Idealmente a janela de pesquisa, utilizada para localizar blocos semelhantes entre duas vistas, deveria ser determinada de acordo com a disparidade existente entre o conteúdo de cada vista. No entanto, como a disparidade depende de múltiplos fatores, nomeadamente, da distância dos objetos no conteúdo da imagem às câmaras, distância e tipo de alinhamento entre os eixos óticos das câmaras, torna-se difícil conseguir determinar um valor

fixo para o tamanho da janela de pesquisa. Com a realização de testes para determinar o tamanho da janela de pesquisa mais favorável ao desempenho dos métodos, foi possível concluir que, de uma forma geral, existe um valor a partir do qual a janela de pesquisa fornece sempre o melhor resultado.

Apesar de ambos os métodos de melhoria da vista auxiliar utilizarem a mesma funcionalidade para a estimação de disparidade através da pesquisa de bloco semelhante, o método baseado na malha triangular executa este processo de uma forma particular. A obtenção de pontos característicos da imagem para a localização dos vértices da malha a construir sobre o conteúdo desta, permite obter as zonas onde o processo de pesquisa de bloco, com recurso a algoritmos de medição de similaridade, funciona melhor. Esta situação beneficia o desenvolvimento da posterior adaptação de conteúdo relativo aos detalhes da vista base para a vista auxiliar.

No processo de adição de detalhes, a adição direta dos detalhes da vista base adaptados à vista auxiliar resulta no aparecimento de ruído em zonas onde o método de melhoria não tem um bom desempenho. Nomeadamente, zonas da imagem com conteúdo homogéneo, ou seja, com um baixo nível de detalhe. A utilização de um método de adição de detalhes, baseado na máscara de contornos da imagem, permite adicionar detalhes apenas em zonas não homogéneas, beneficiando o desempenho do método de melhoria da imagem auxiliar.

Com a verificação da influência dos detalhes (*sharpness*) da imagem nos métodos de melhoria da vista auxiliar em comparação, foi sentida a dificuldade de fazer corresponder um valor objetivo do nível de detalhe das imagens a processar, com os resultados obtidos. Como trabalho futuro poderá ser desenvolvido um método para analisar o detalhe de cada área da imagem de forma objetiva, de modo a determinar a correspondência entre o detalhe da imagem e os resultados obtidos em cada método.

A maior parte das sequências estéreo codificadas de forma assimétrica, com resoluções diferentes, melhoram o seu desempenho com os métodos propostos. Apenas as sequências de baixa resolução e sequências com oclusões significativas de conteúdo entre ambas as vistas tiveram um desempenho fraco ou negativo, não justificando a utilização de processamento adicional com os métodos desenvolvidos.

Comparando o desempenho dos métodos propostos, foram obtidos ganhos até 1.55 dB, em termos de *Delta Bjøntegaard* PSNR (BD-PSNR) com o método baseado na malha triangular em conjunto com o método da adição baseada na máscara de contornos, relativamente a 1.38 BD-PSNR obtido com o método baseado na pesquisa de blocos e adição baseada na máscara de contornos, no caso da vista auxiliar subamostrada simultaneamente na horizontal e vertical.

Desenvolvimentos futuros deverão incluir testes experimentais para verificar possíveis diferenças na perceção de profundidade nas sequências melhoradas, através de testes sub-

jetivos. Testes adicionais poderão incluir subamostragem da vista auxiliar por outros fatores além de 2. A subamostragem da imagem auxiliar poderá ser substituída, ou combinada, com a utilização de diferentes parâmetros de quantificação (QP) para cada vista.

Bibliografia

- [1] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC Standard," *Proceedings of the IEEE*, vol. 99, 2011.
- [2] L. B. Stelmach, W. J. Tam, D. V. Meegan, A. Vincent, and P. Corriveau, "Human perception of mismatched stereoscopic 3D inputs," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Vancouver, US, Sept. 2000.
- [3] M. G. Perkins, "Data compression of stereopairs," *IEEE Transaction on Communications*, vol. 2, Apr. 1992.
- [4] I. K. Kim, K. McCann, K. Sugimoto, B. Bros, W. J. Han, and G. Sullivan, "High efficiency video coding (HEVC) test model 12 HM12 encoder description," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 14th Meeting*, Vienna, Austria, Jul. 2013.
- [5] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, Dec. 2012.
- [6] A. Aksay, C. Bilen, E. Kurutepe, T. Ozcelebi, G. B. Akar, M. R. Civanlar, and A. M. Tekalp, "Temporal and spatial scaling for stereoscopic video compression," in *Proceedings of 14th European Signal Processing Conference*, Florence, Italy, Sept. 2006.
- [7] C. Fehn, P. Kauff, S. Cho, H. Kwon, N. Hur, and J. Kim, "Asymmetric coding of stereoscopic video for transmission over T-DMB," in *3DTV Conference, 2007*, Kos Island, Greece, May 2007.
- [8] Y. Chen, Y.-K. Wang, M. Gabbouj, and M. M. Hannuksela, "Regionally adaptive filtering for asymmetric stereoscopic video coding," in *Proc. of the IEEE International Symposium on Circuits and Systems*, Taipei, Taiwan, May 2009.

-
- [9] F. Shao, G. Jiang, X. Wang, M. Yu, and K. Chen, "Stereoscopic video coding with asymmetric luminance and chrominance qualities," *IEEE Transactions on Consumer Electronics*, vol. 56, Nov. 2010.
- [10] S. Valizadeh, M. Azimi, and P. Nasiopoulos, "Bitrate reduction in asymmetric stereoscopic video with low-pass filtered slices," in *IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, US, Jan. 2012.
- [11] M. Azimi, S. Valizadeh, L. Xiaokang, L. E. Coria, and P. Nasiopoulos, "Subjective study on asymmetric stereoscopic video with low-pass filtered slices," in *International Conference on Computing, Networking and Communications (ICNC)*, Maui, Hawaii, Jan. 2012.
- [12] J. Tian, L. Chen, and Z. Liu, "Asymmetric stereoscopic image resolution enhancement," in *First Asian Conference on Pattern Recognition (ACPR)*, Beijing, China, Nov. 2011.
- [13] F. Brandi, R. de Queiroz, and D. Mukherjee, "Super-resolution of video using key frames and motion estimation," in *15th (IEEE) International Conference on Image Processing, 2008 (ICIP 2008)*, San Diego, US, Oct. 2008.
- [14] C. Dórea, D. Garcia, and R. de Queiroz, "Super-resolução para síntese de vistas em ambientes de resolução mista," in *XXXI Simpósio Brasileiro de Telecomunicações, 2013 (SBrT 2013)*, Fortaleza, Brazil, Sept. 2013.
- [15] ISO IEC JTC1 SC29 WG11 and ITU-T Q6 16, "Joint call for proposals on video compression technology," in *WG11 Doc. N11113 and Q6/16 Doc. VCEG-AM91*, Kyoto, Japan, Jan. 2010.
- [16] JCTVC, "Encoder-side description of test model under consideration," in *JCTVC B204*, Geneva, Switzerland, Jul. 2010.
- [17] K. McCann, B. Bross, S. ichi Sekiguchi, and W. J. Han, "High efficiency video coding (HEVC) test model 1 (HM 1) encoder description," in *JCTVC C402*, Guangzhou, China, Oct. 2010.
- [18] P. Bordes, G. Clare, F. Henry, M. Raulet, and J. Viéron, "An overview of the emerging HEVC Standard," Jul. 2012.
- [19] M. Pourazad, C. Dautre, M. Azimi, and P. Nasiopoulos, "HEVC: The new gold standard for video compression: How does HEVC compare with H.264/AVC?" *Consumer Electronics Magazine, IEEE*, vol. 1, Jul. 2012.

- [20] I.-K. Kim, S. Lee, M.-S. Cheon, T. Lee, and J. Park, "Coding efficiency improvement of HEVC using asymmetric motion partitioning," in *Broadband Multimedia Systems and Broadcasting (BMSB), 2012 IEEE International Symposium on*, Seoul, Jun. 2012.
- [21] I.-K. Kim, W.-J. Han, J. Park, and X. Zheng, "CE2: Test results of asymmetric motion partition (amp)," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 10th Meeting*, Torino, Italy, Jul. 2011.
- [22] Y.-S. Ho *et al.*, *Advanced Video Coding for Next-Generation Multimedia Services*, 1st ed. InTech, 2012.
- [23] V. Seferidis and D. V. Papadimitriou, "Improved disparity estimation in stereoscopic television," *Electronics Letters, Issue 9*, vol. 29, 1993.
- [24] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-8, Issue 6*, Nov. 1986.
- [25] J. Shi and C. Tomasi, "Good features to track," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, Seattle, US, Jun. 1994.
- [26] C. Harris and M. Stephens, "A combined corner and edge detector," *Proceedings of the Fourth Alvey Vision Conference*, 1988.
- [27] B. N. Delaunay, "Sur la sphère vide," *Bulletin de L'Académie des Sciences de L'URSS, Classe des Sciences Mathématiques et Naturelles*, 1934.
- [28] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd ed., ser. Wiley Series in Probability and Statistics. Wiley, 2000.
- [29] [Acedido em 11-Junho-2014]. [Online]. Available: <http://mathworld.wolfram.com/AffineTransformation.html>
- [30] [Acedido em 11-Junho-2014]. [Online]. Available: http://docs.opencv.org/doc/tutorials/imgproc/imgtrans/warp_affine/warp_affine.html
- [31] J. R. Ohm, G. J. Sullivan, H. Schwarz, and T. K. Tan, "Comparison of the coding efficiency of video coding standards - including high efficiency video coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, Dec. 2012.

- [32] F. Bossen, "Common test conditions and software reference configurations," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 10th Meeting*, Stockholm, Sweden, Jul. 2012.
- [33] J. Zhang, R. Li, H. Li, D. Rusanovskyy, and M. M. Hannuksela, "Ghost Town Fly 3DV sequence for purposes of 3DV standardization," in *MPEG 2011/M20027*, Geneve, Switzerland, Mar. 2011.
- [34] M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, and K. Wegner, "Poznan multiview video test sequences and camera parameters," in *International Organization for Standardisation ISO/IEC JTC1/SC29/WG11 MPEG/17050*, Xian, China, Oct. 2009.
- [35] M. Tanimoto, T. Fugi, and N. Fukushima, "1D parallel test sequences for MPEG-FTV," in *International Organization for Standardisation ISO/IEC JTC1/SC29/WG11 M15378*, Archamps, France, Apr. 2008.
- [36] I. Feldmann, M. Mueller, F. Zilly, R. Tanger, K. Mueller, A. Smolic, P. Kauff, and T. Wiegand, "HHI test material for 3D video," in *International Organization for Standardisation ISO/IEC JTC1/SC29/WG11 MPEG/15413*, Archamps, France, Apr. 2008.
- [37] A. Vetro, M. McGuire, W. Matusik, A. B. J. Lee, and H. Pfister, "Multiview video test sequences from merl," in *International Organization for Standardisation ISO/IEC JTC1/SC29/WG11 MPEG/m12077*, Busan, Korea, Apr. 2005.
- [38] G. Jorge, S. M. M. Faria, and N. M. M. Rodrigues, "Affine transform inter-view enhancement in spatial asymmetric stereo video decoding," in *9th Conference on Telecommunications, Conftele 2013*, Castelo Branco, Portugal, May 2013.
- [39] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," in *VCEG Contribution VCEG-M33*, Austin, US, Apr. 2001.
- [40] [Acedido em 11-Junho-2014]. [Online]. Available: <http://gcc.gnu.org>
- [41] [Acedido em 11-Junho-2014]. [Online]. Available: <http://opencv.org>
- [42] [Acedido em 11-Junho-2014]. [Online]. Available: <http://www.hhi.fraunhofer.de/de/kompetenzfelder/image-processing/research-groups/image-video-coding/svc-extension-of-h264avc/jsvm-reference-software.html>

Apêndice A

Testes de parâmetros

A.1 Tamanho do *kernel* para o método da malha triangular

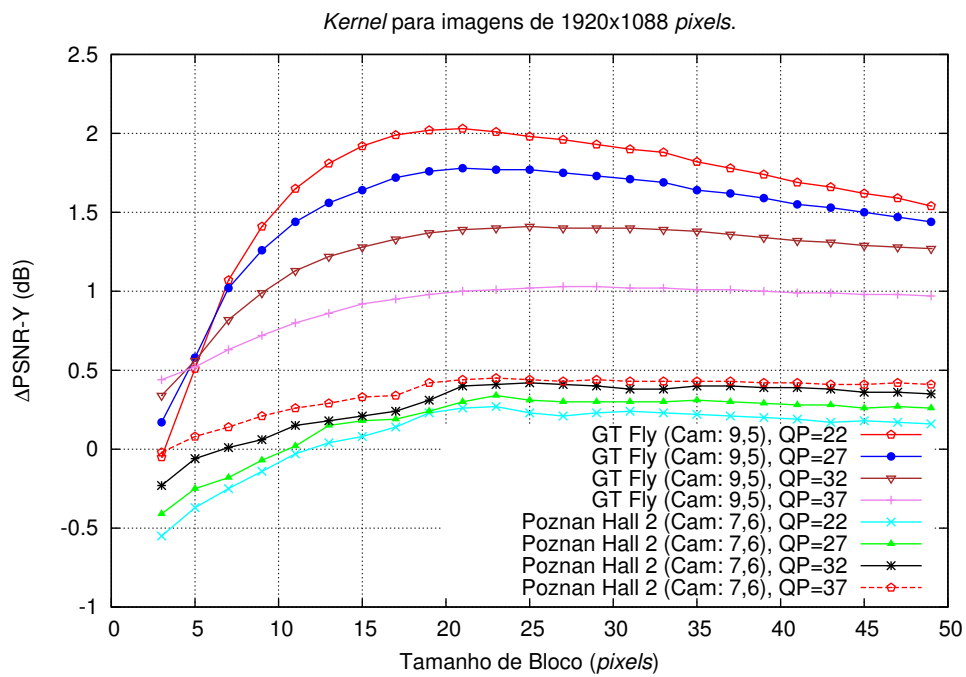


Figura A.1: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do *kernel*, com a vista auxiliar subamostrada na horizontal (H).

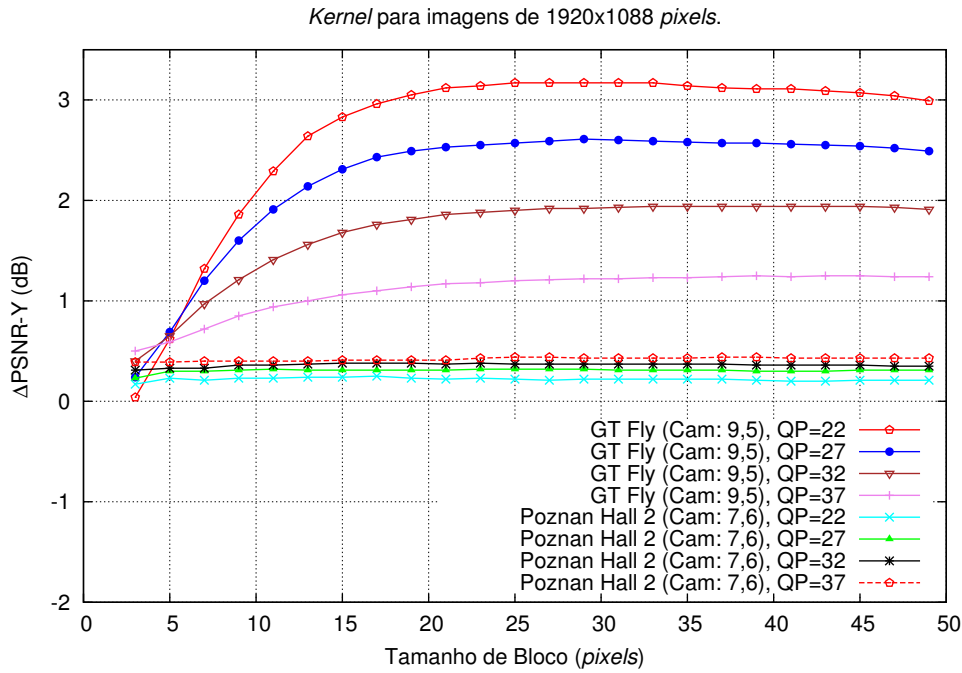


Figura A.2: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do *kernel*, com a vista auxiliar subamostrada na vertical (V).

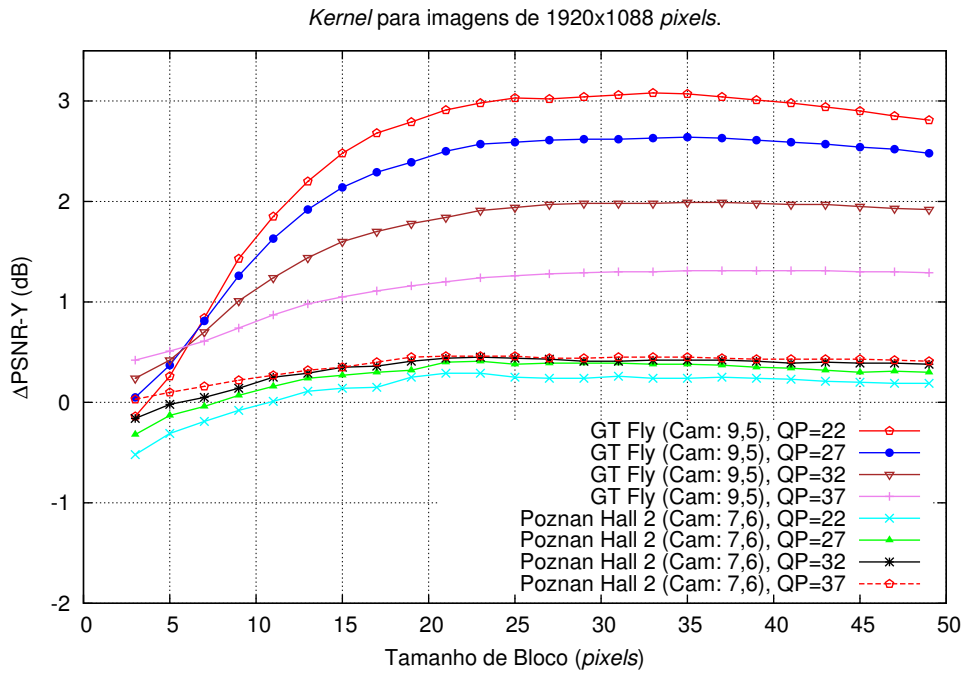


Figura A.3: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do *kernel*, com a vista auxiliar subamostrada na horizontal e vertical (H/V).

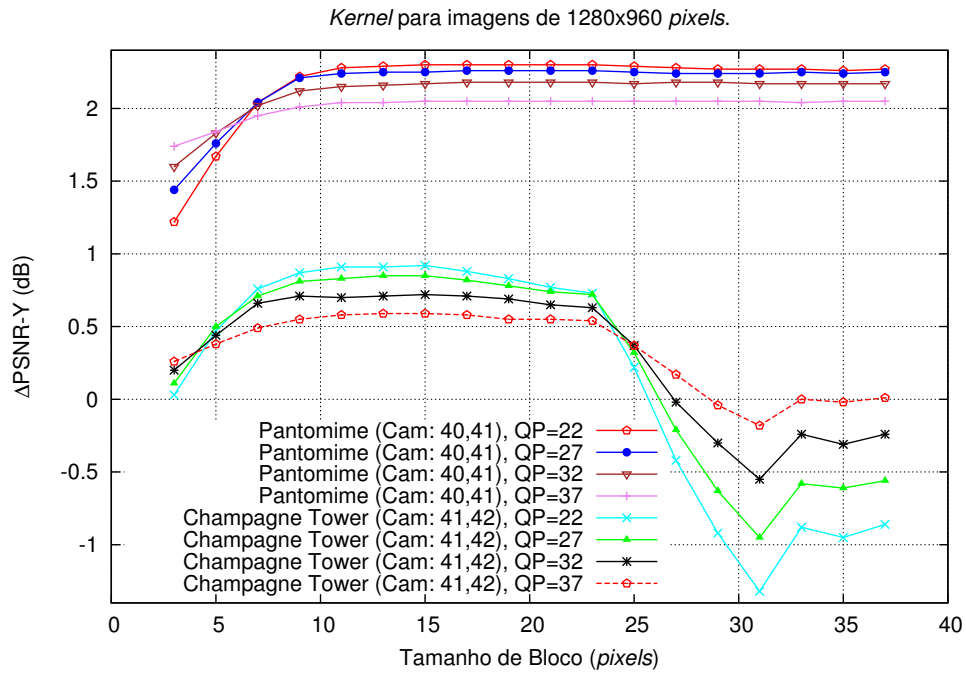


Figura A.4: Diferença de PSNR das sequências *Pantomime* e *Champagne Tower* em função da variação do tamanho do *kernel*, com a vista auxiliar H.

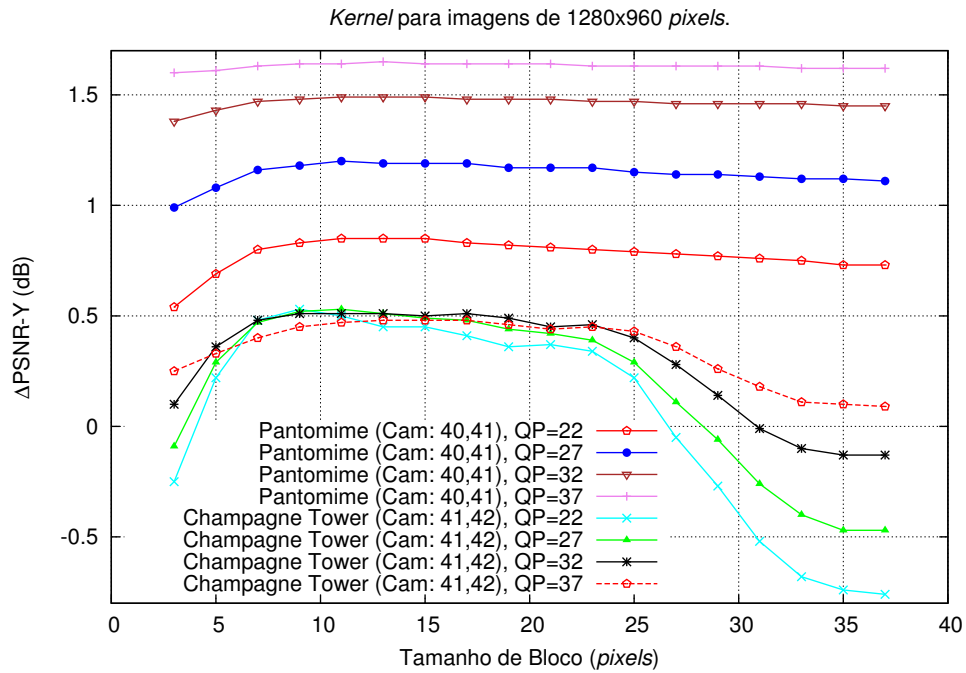


Figura A.5: Diferença de PSNR das sequências *Pantomime* e *Champagne Tower* em função da variação do tamanho do *kernel*, com a vista auxiliar V.

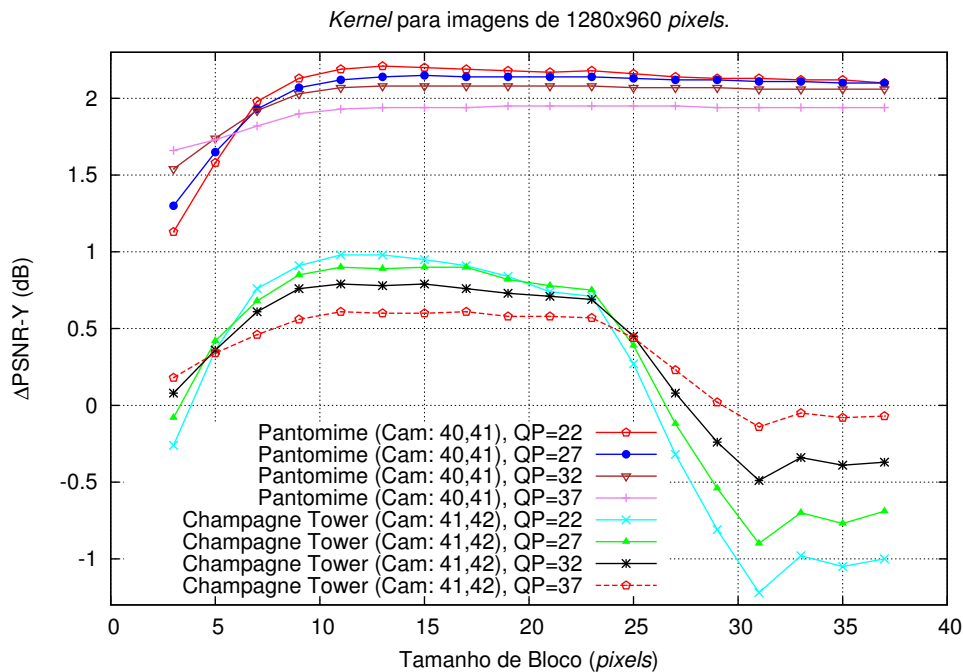


Figura A.6: Diferença de PSNR das seqüências *Pantomime* e *Champagne Tower* em função da variação do tamanho do *kernel*, com a vista auxiliar H/V.

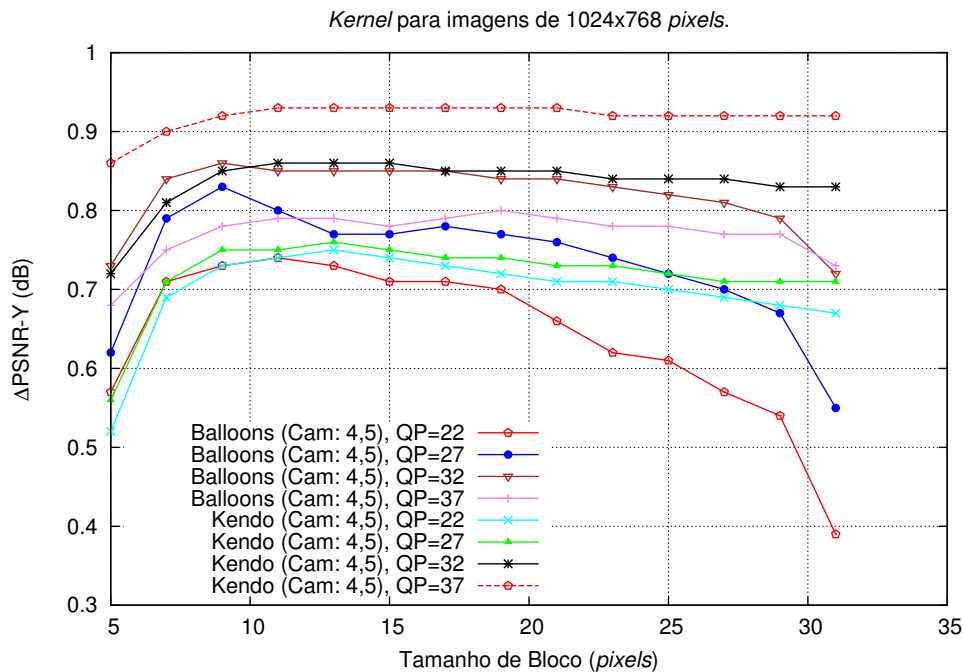


Figura A.7: Diferença de PSNR das seqüências *Balloons* e *Kendo* em função da variação do tamanho do *kernel*, com a vista auxiliar H.

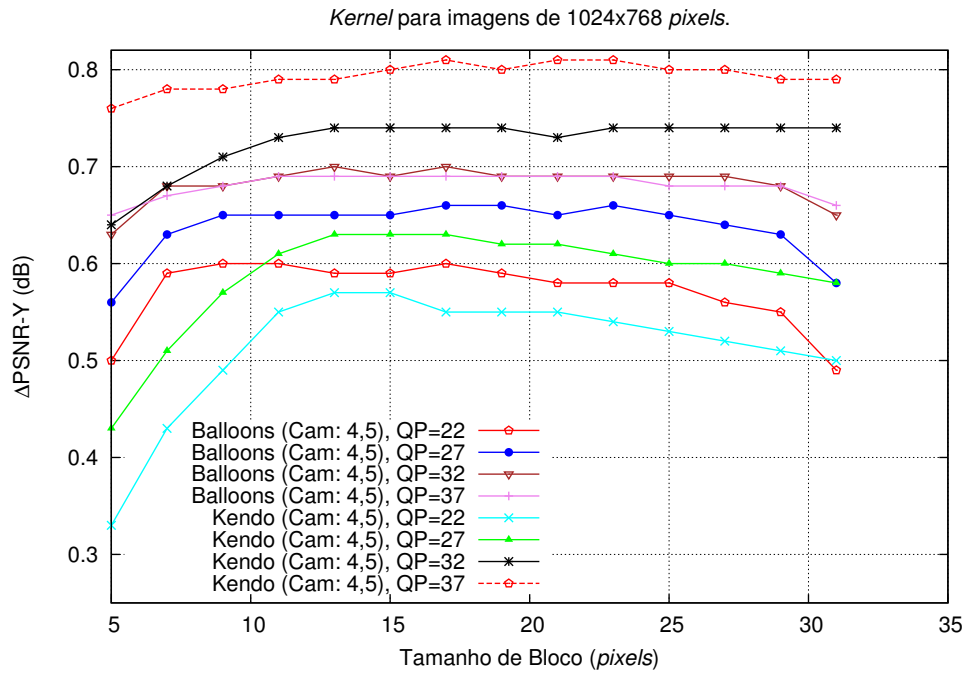


Figura A.8: Diferença de PSNR das sequências *Balloons* e *Kendo* em função da variação do tamanho do *kernel*, com a vista auxiliar V.

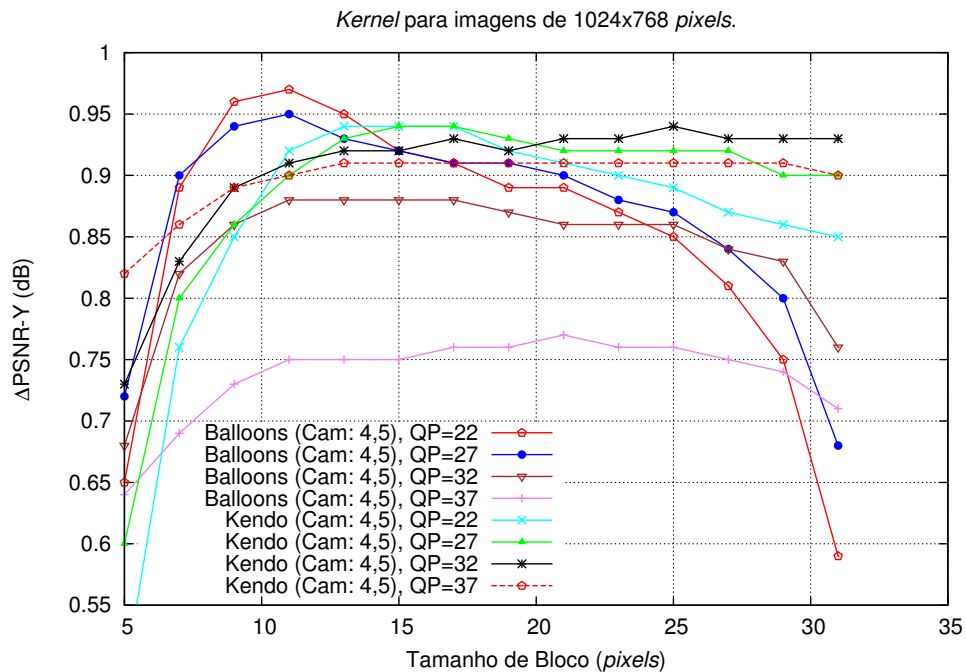


Figura A.9: Diferença de PSNR das sequências *Balloons* e *Kendo* em função da variação do tamanho do *kernel*, com a vista auxiliar H/V.

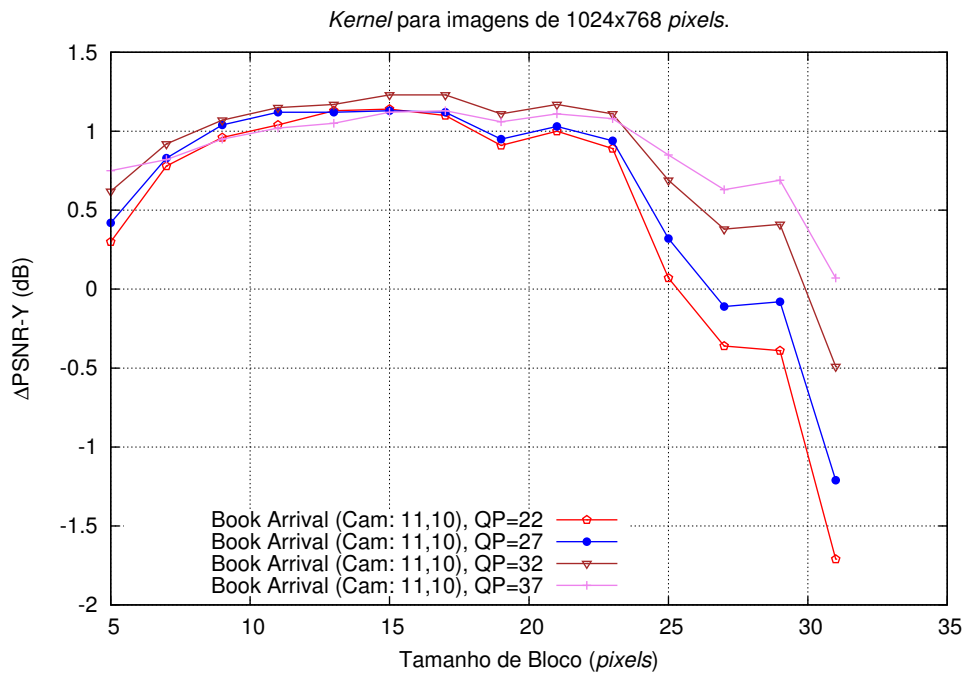


Figura A.10: Diferença de PSNR da sequência *Book Arrival* em função da variação do tamanho do *kernel*, com a vista auxiliar H.

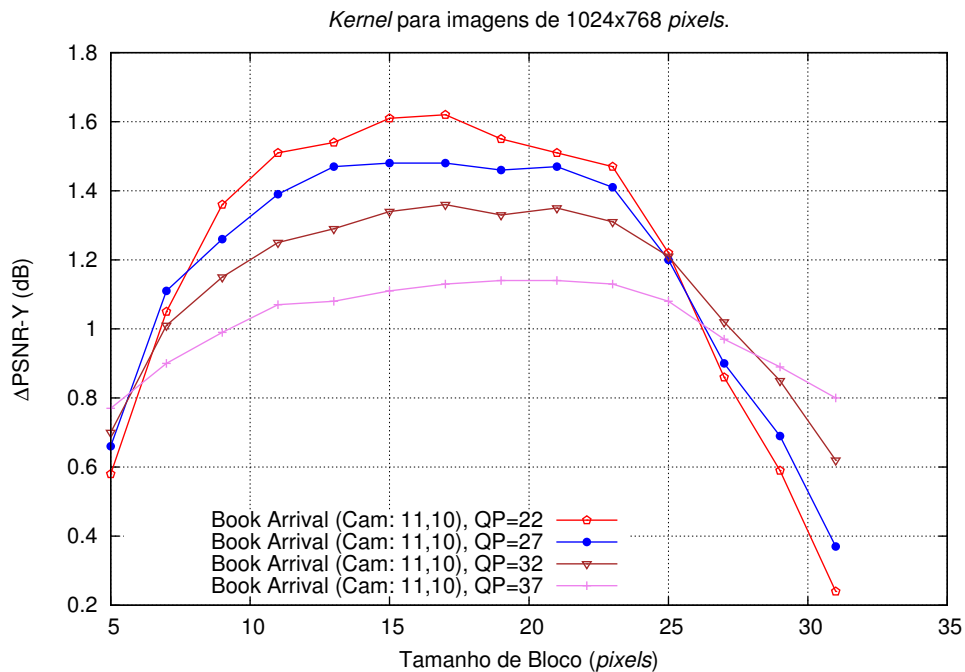


Figura A.11: Diferença de PSNR da sequência *Book Arrival* em função da variação do tamanho do *kernel*, com a vista auxiliar V.

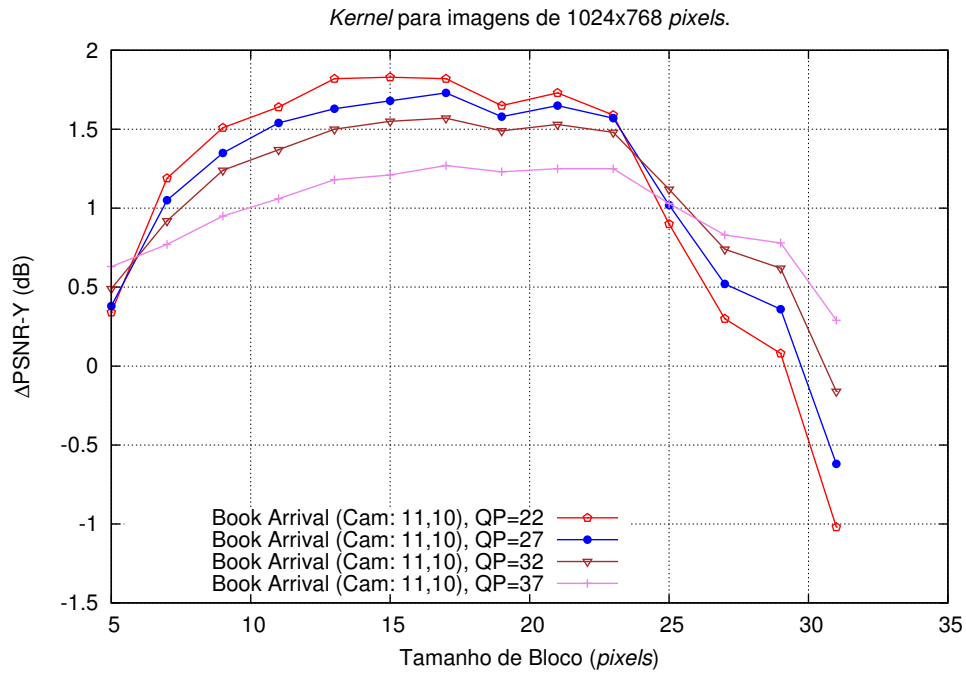


Figura A.12: Diferença de PSNR da sequência *Book Arrival* em função da variação do tamanho do *kernel*, com a vista auxiliar H/V.

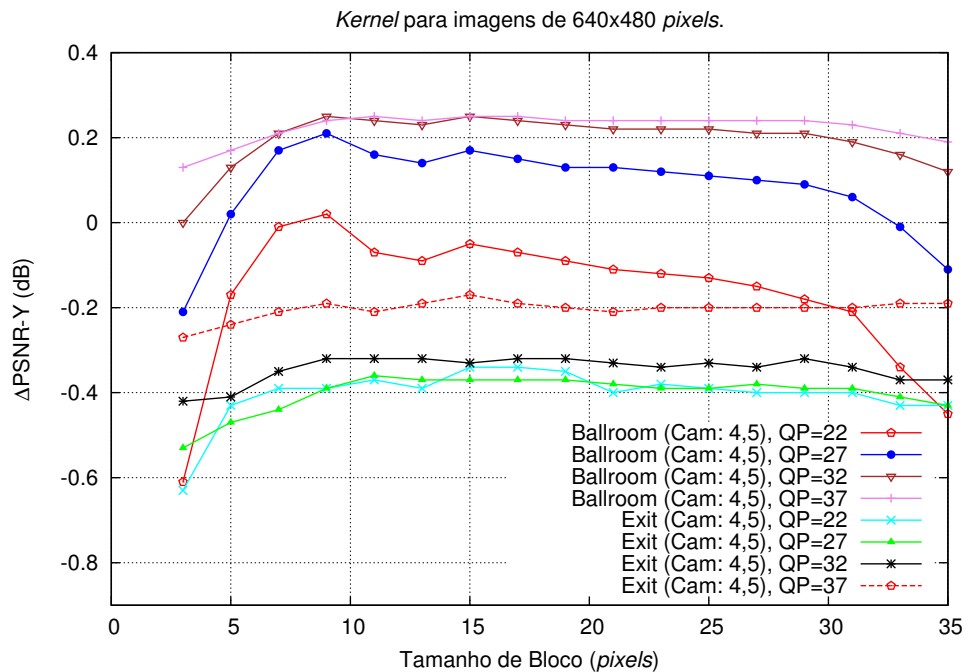


Figura A.13: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho do *kernel*, com a vista auxiliar H.

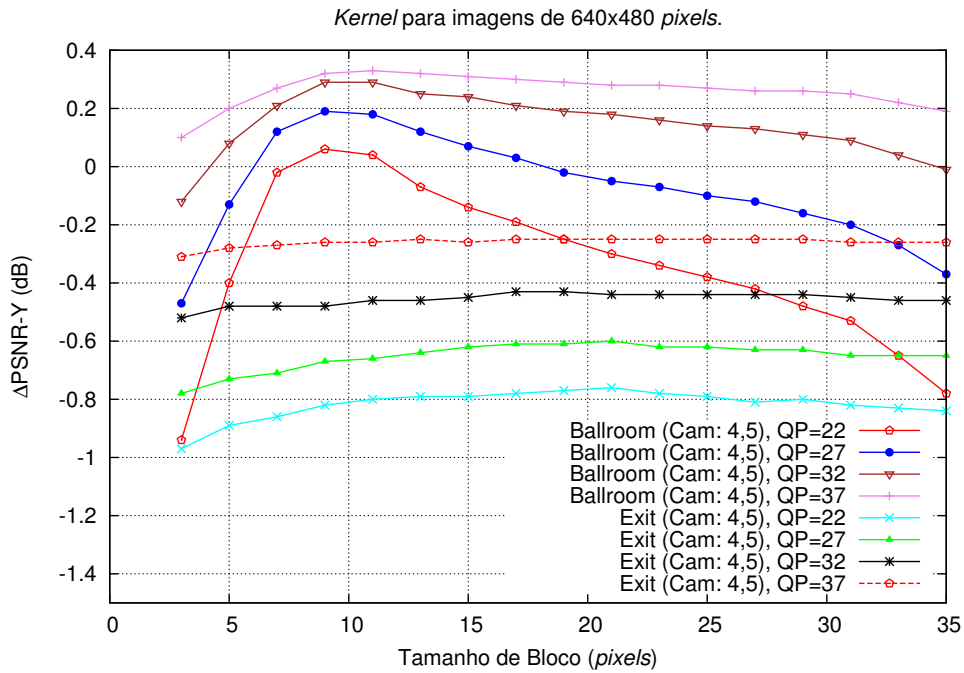


Figura A.14: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho do *kernel*, com a vista auxiliar V.

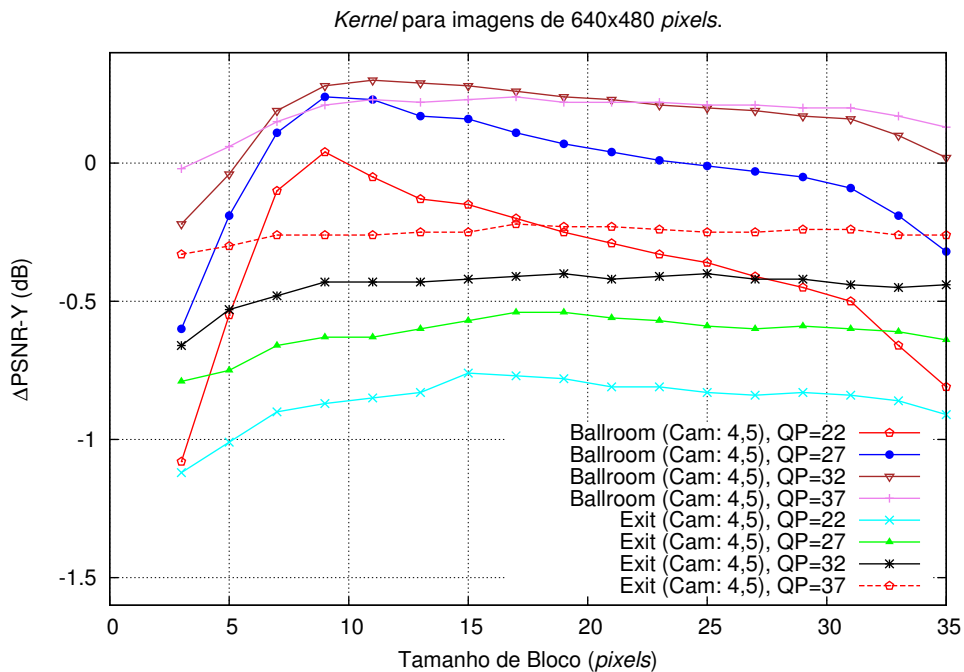


Figura A.15: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho do *kernel*, com a vista auxiliar H/V.

A.2 Largura da janela de pesquisa para o método baseado na malha triangular

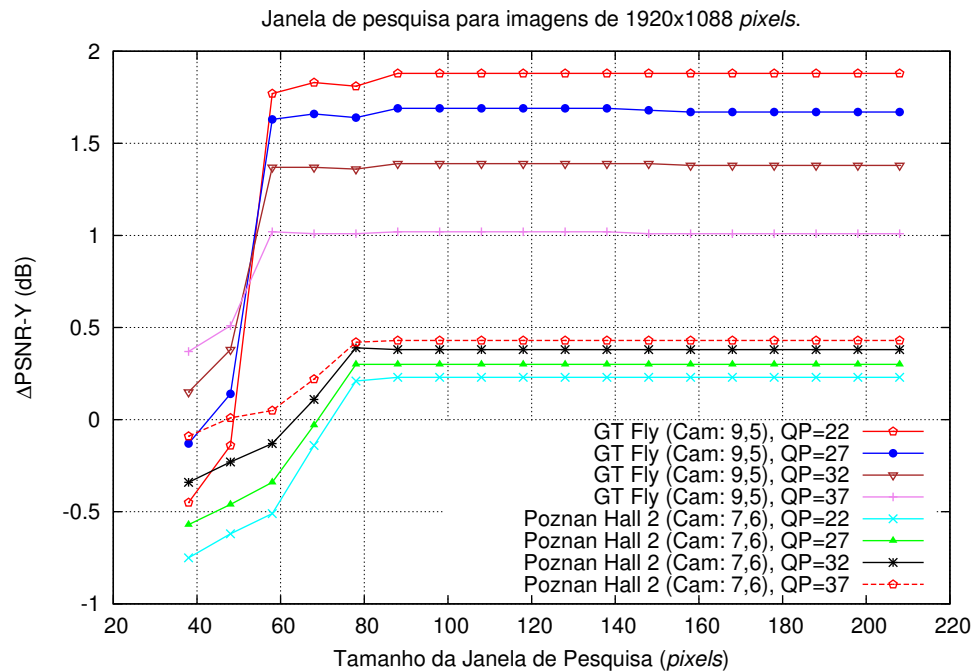


Figura A.16: Diferença de PSNR das seqüências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

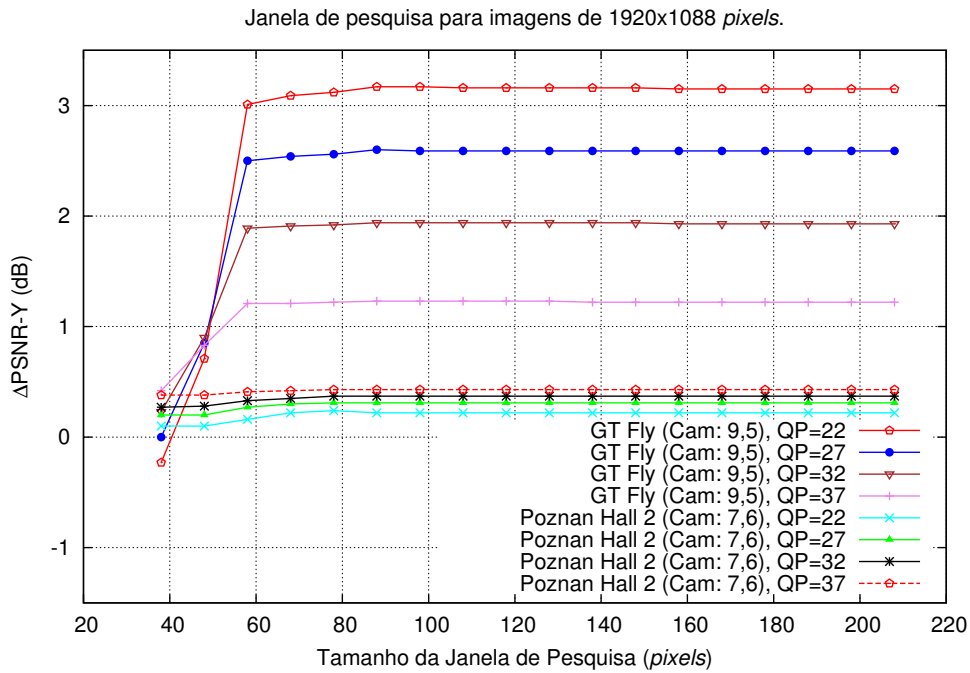


Figura A.17: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

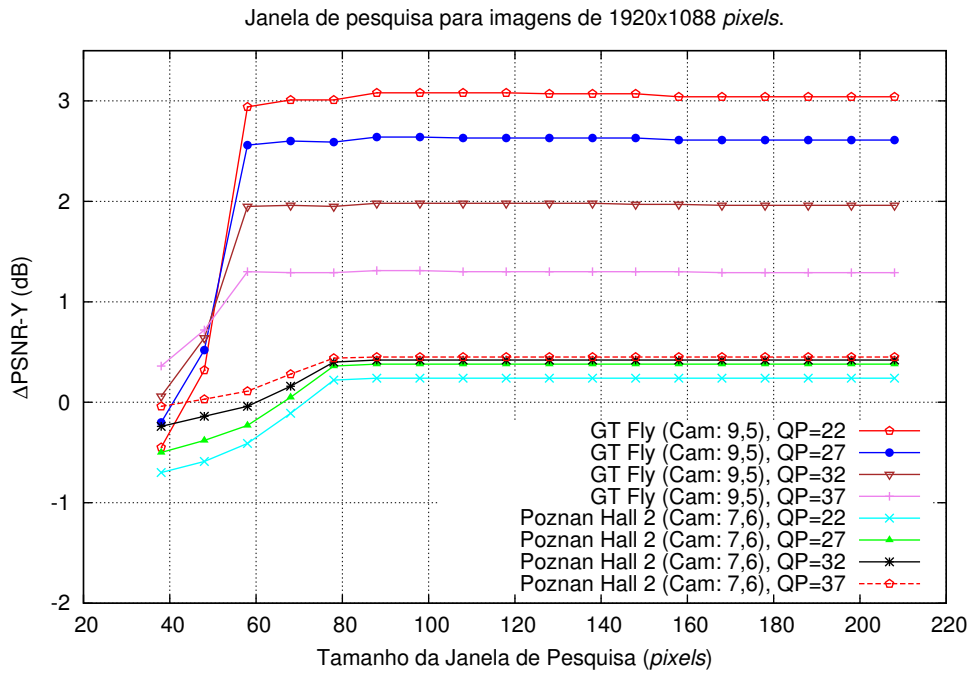


Figura A.18: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

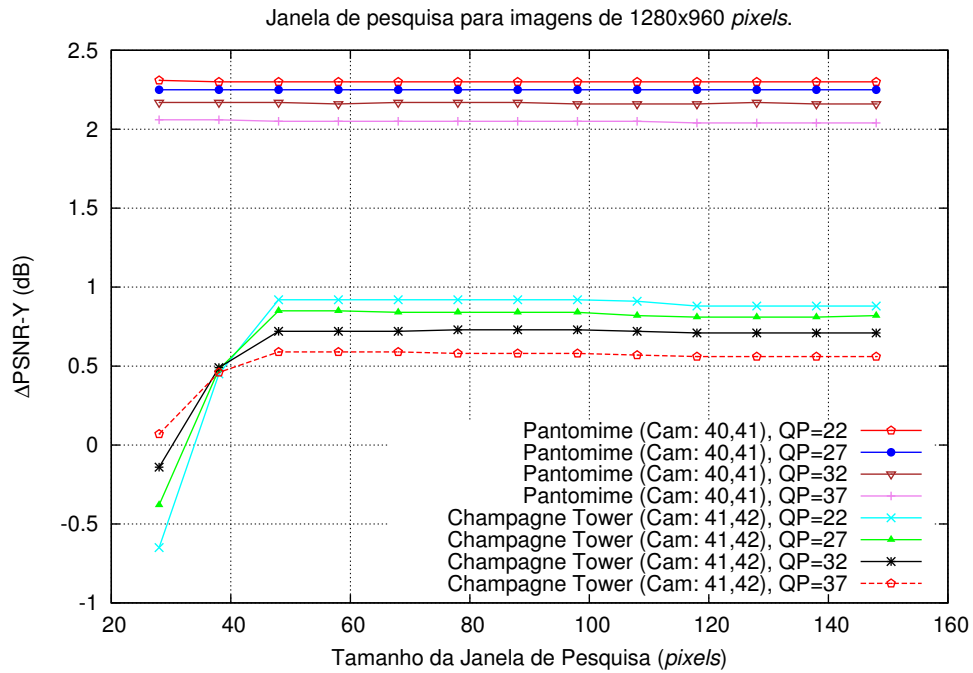


Figura A.19: Diferença de PSNR das sequências *Pantomime* e *Champagne Tower* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

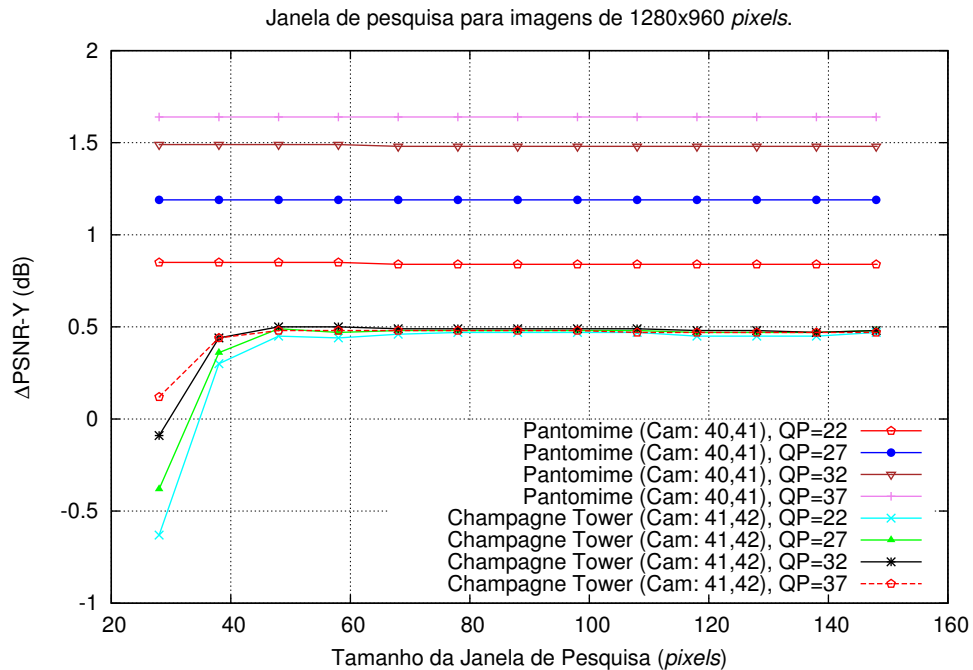


Figura A.20: Diferença de PSNR das sequências *Pantomime* e *Champagne Tower* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

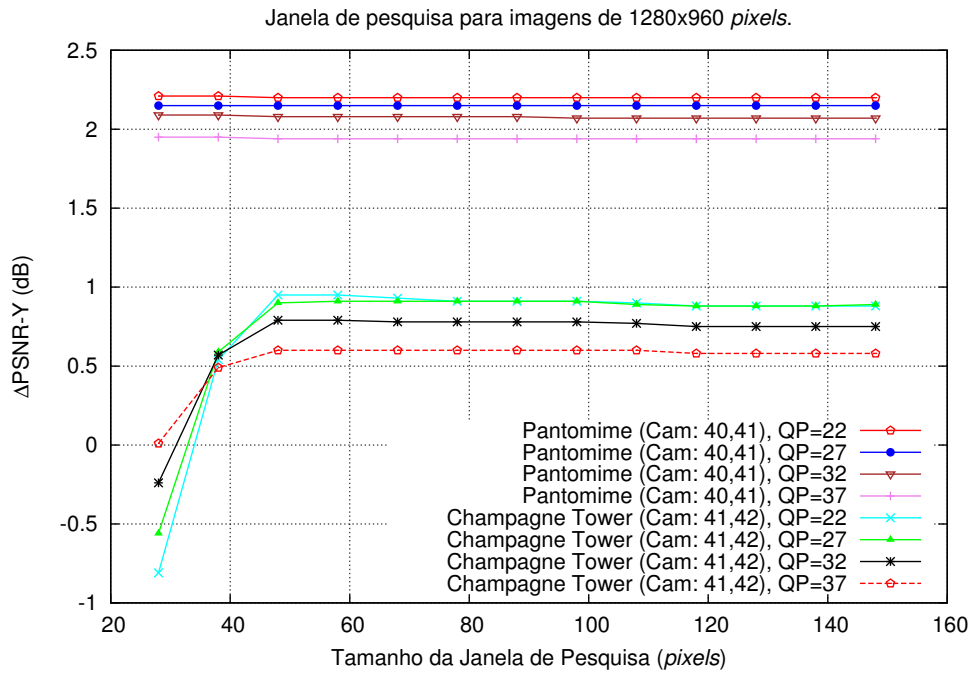


Figura A.21: Diferença de PSNR das seqüências *Pantomime* e *Champagne Tower* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

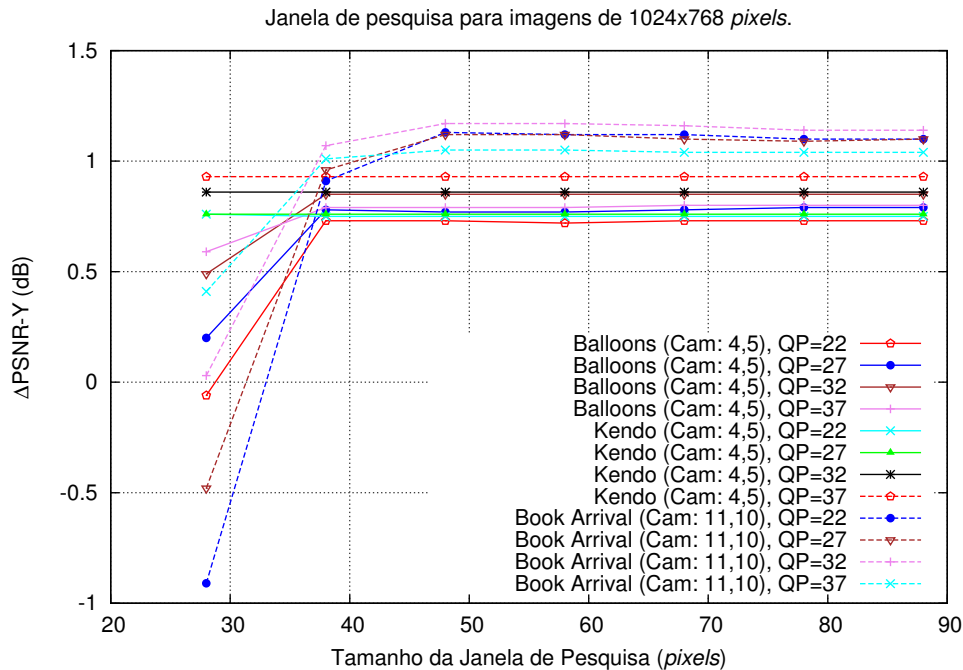


Figura A.22: Diferença de PSNR das seqüências *Balloons* e *Kendo* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

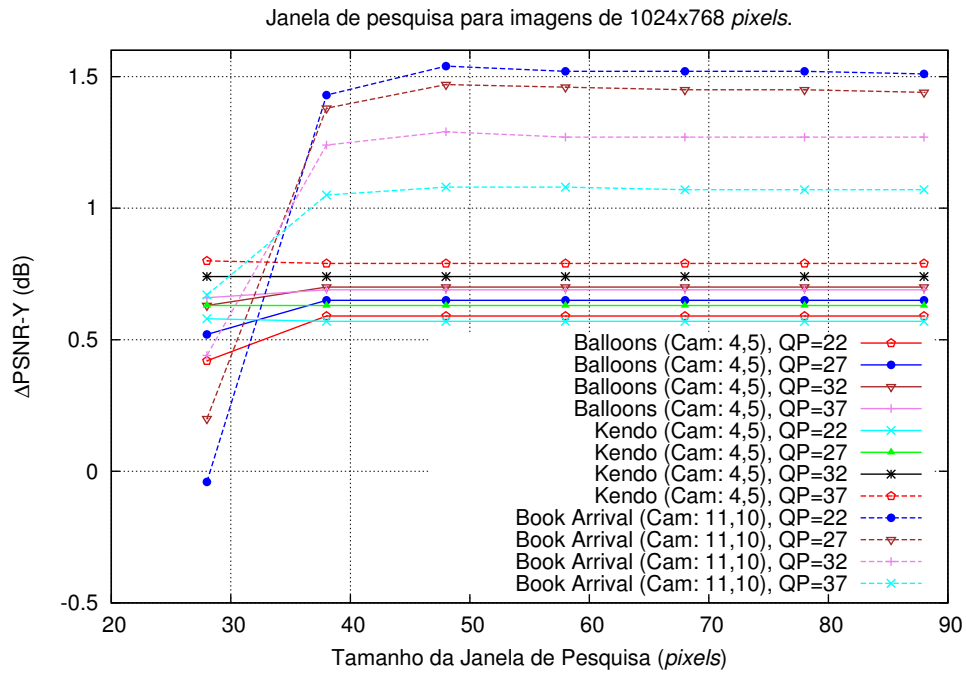


Figura A.23: Diferença de PSNR das sequências *Balloons* e *Kendo* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

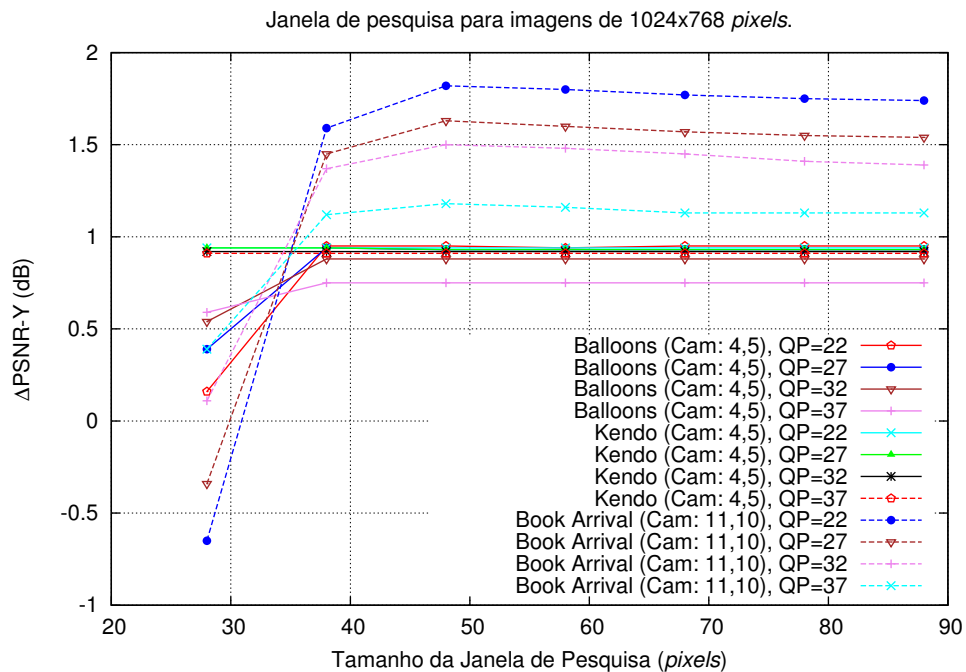


Figura A.24: Diferença de PSNR das sequências *Balloons* e *Kendo* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

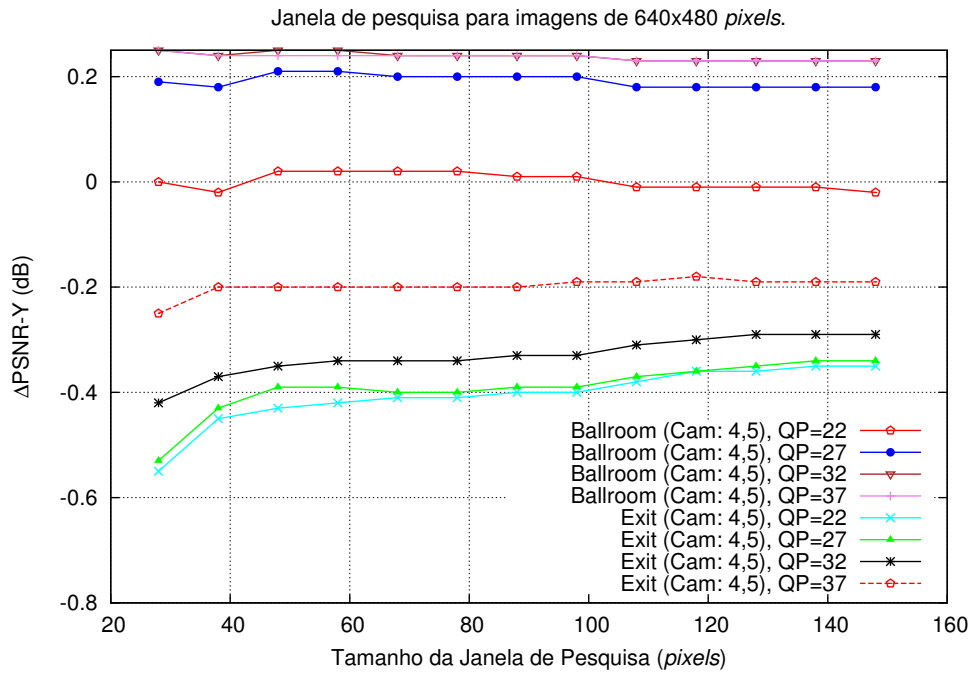


Figura A.25: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

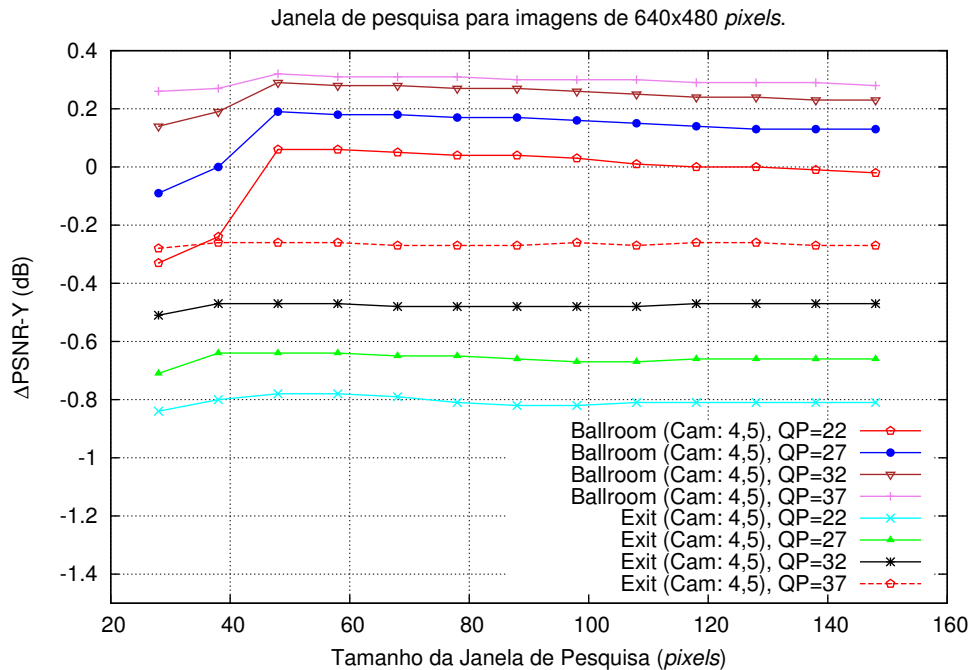


Figura A.26: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

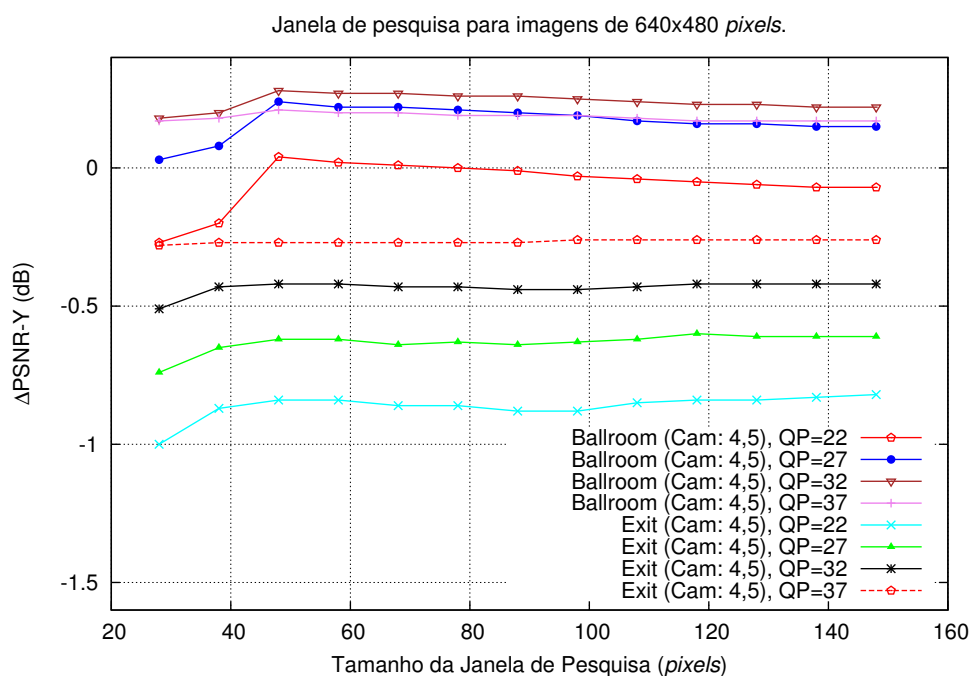


Figura A.27: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

A.3 Tamanho do bloco para o método baseado na pesquisa de blocos

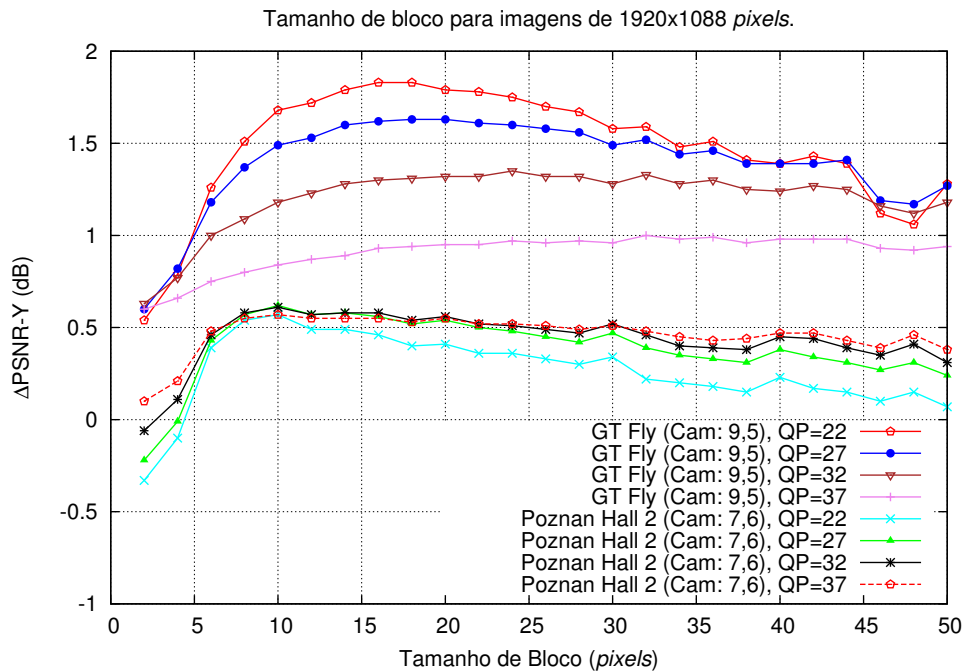


Figura A.28: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do bloco, com a vista auxiliar H.

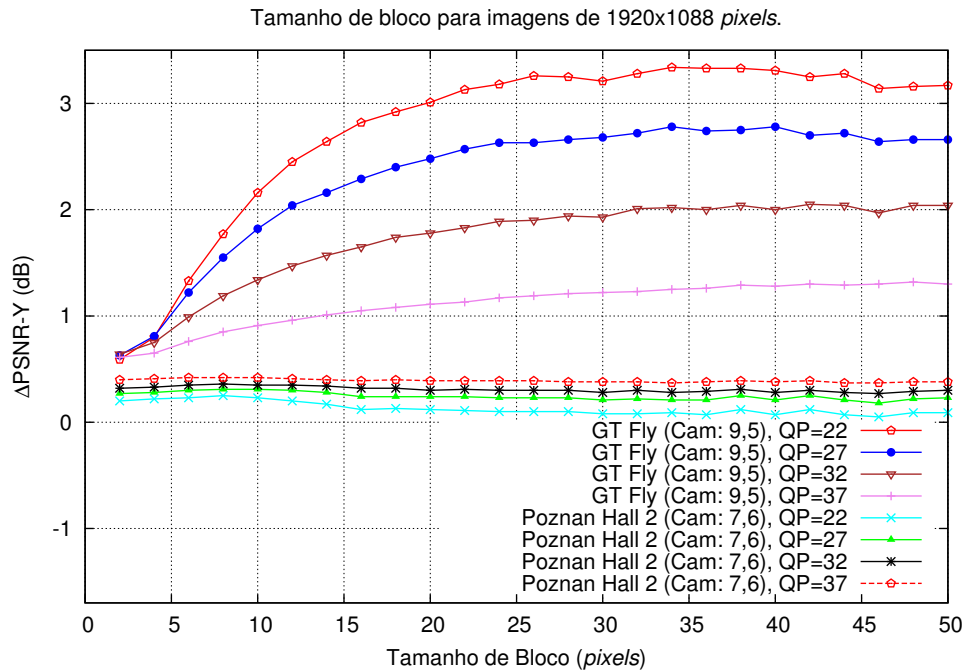


Figura A.29: Diferença de PSNR das seqüências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do bloco, com a vista auxiliar V.

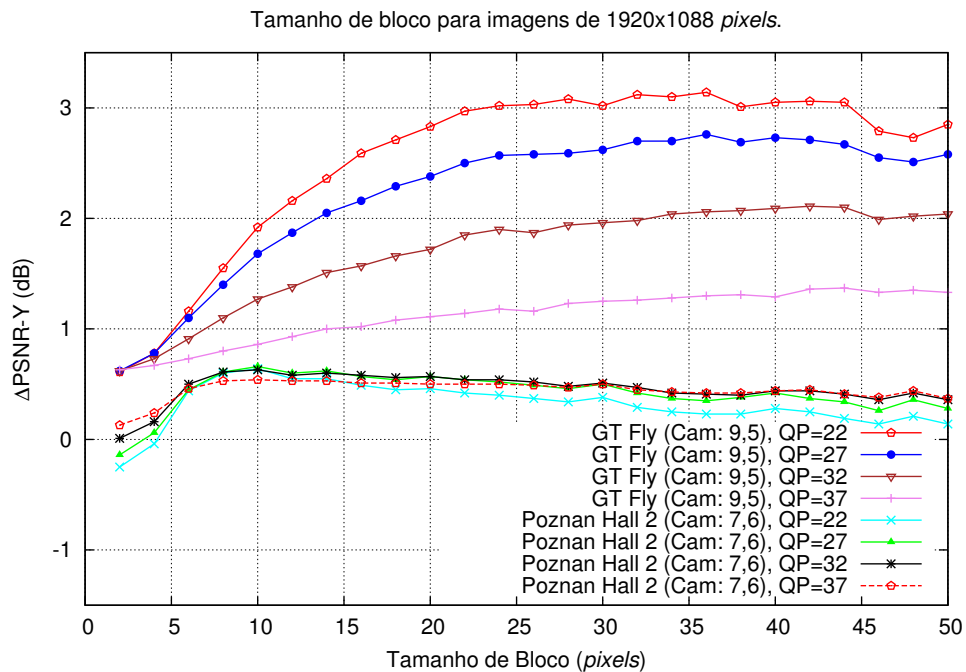


Figura A.30: Diferença de PSNR das seqüências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho do bloco, com a vista auxiliar H/V.

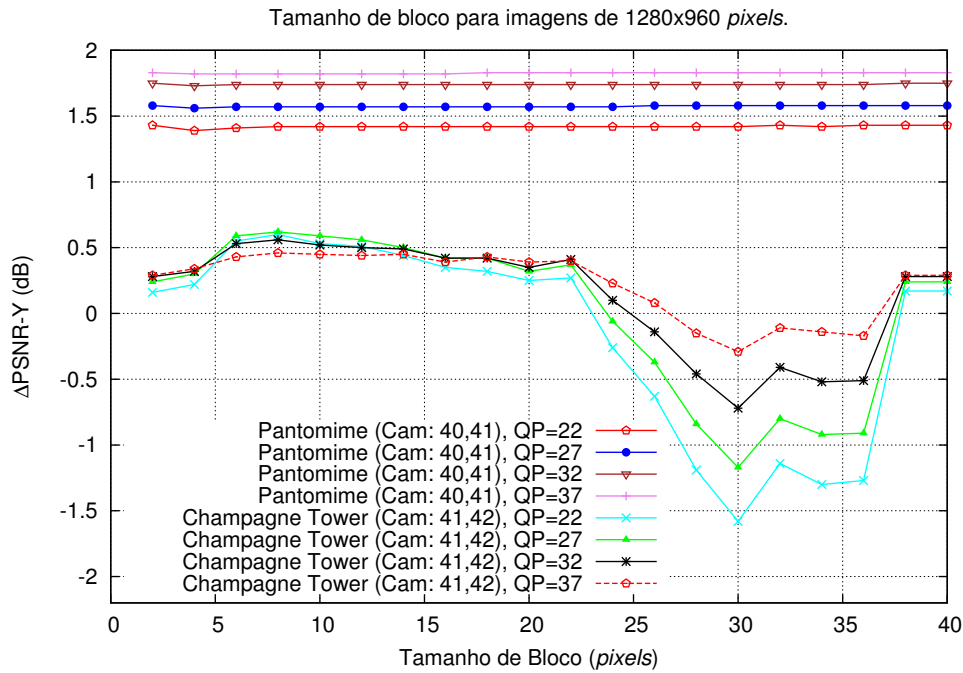


Figura A.31: Diferença de PSNR das seqüências *Pantomime* e *Champagne Tower* em função da variação do tamanho do bloco, com a vista auxiliar H.

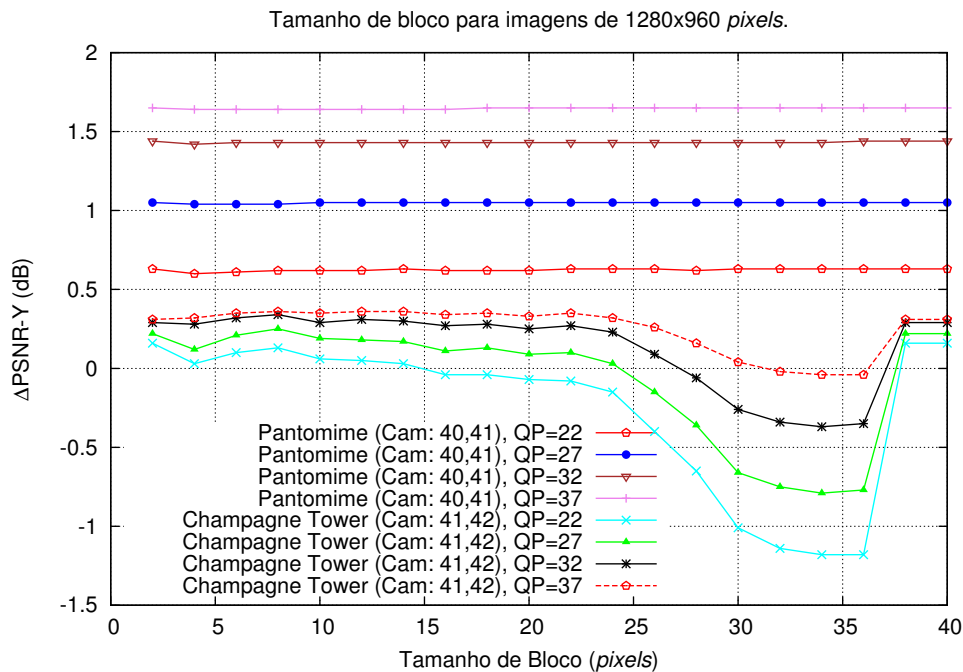


Figura A.32: Diferença de PSNR das seqüências *Pantomime* e *Champagne Tower* em função da variação do tamanho do bloco, com a vista auxiliar V.

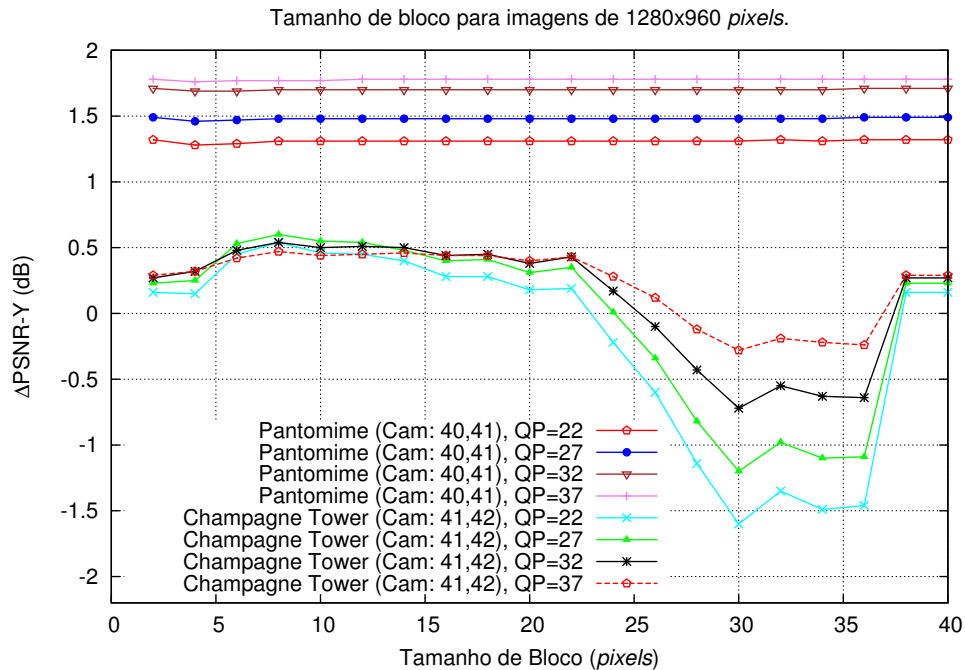


Figura A.33: Diferença de PSNR das sequências *Pantomime* e *Champagne Tower* em função da variação do tamanho do bloco, com a vista auxiliar H/V.

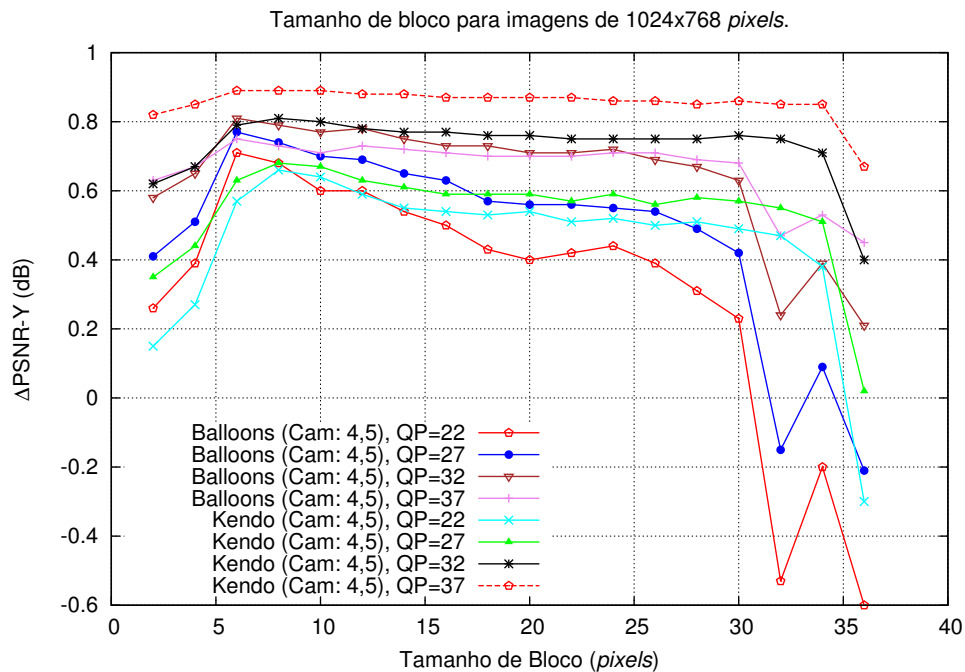


Figura A.34: Diferença de PSNR das sequências *Balloons* e *Kendo* em função da variação do tamanho do bloco, com a vista auxiliar H.

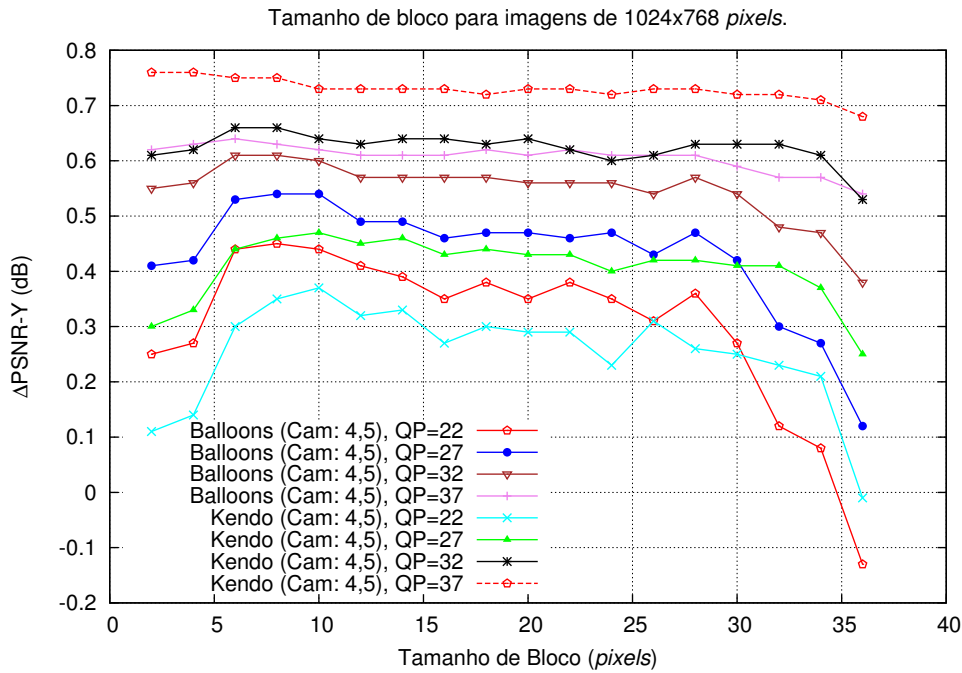


Figura A.35: Diferença de PSNR das seqüências *Balloons* e *Kendo* em função da variação do tamanho do bloco, com a vista auxiliar V.

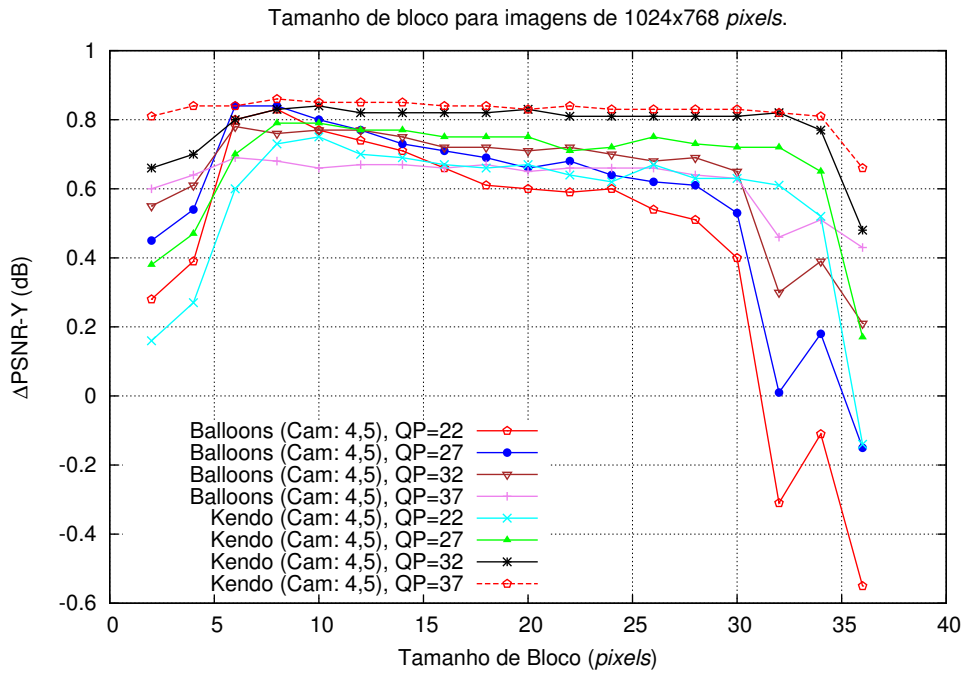


Figura A.36: Diferença de PSNR das seqüências *Balloons* e *Kendo* em função da variação do tamanho do bloco, com a vista auxiliar H/V.

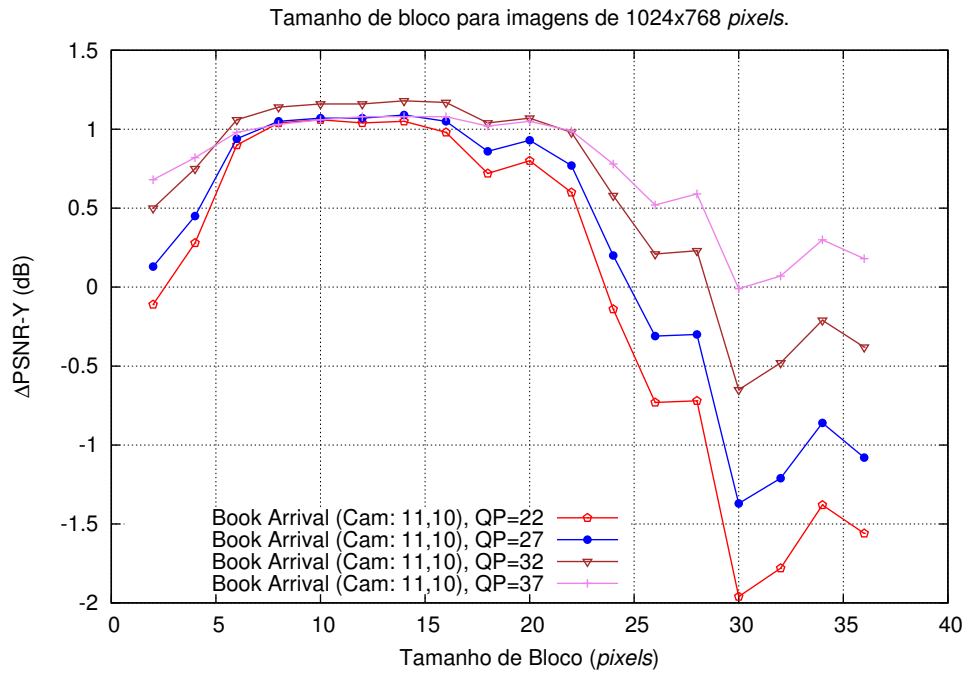


Figura A.37: Diferença de PSNR da sequência *Book Arrival* em função da variação do tamanho do bloco, com a vista auxiliar H.

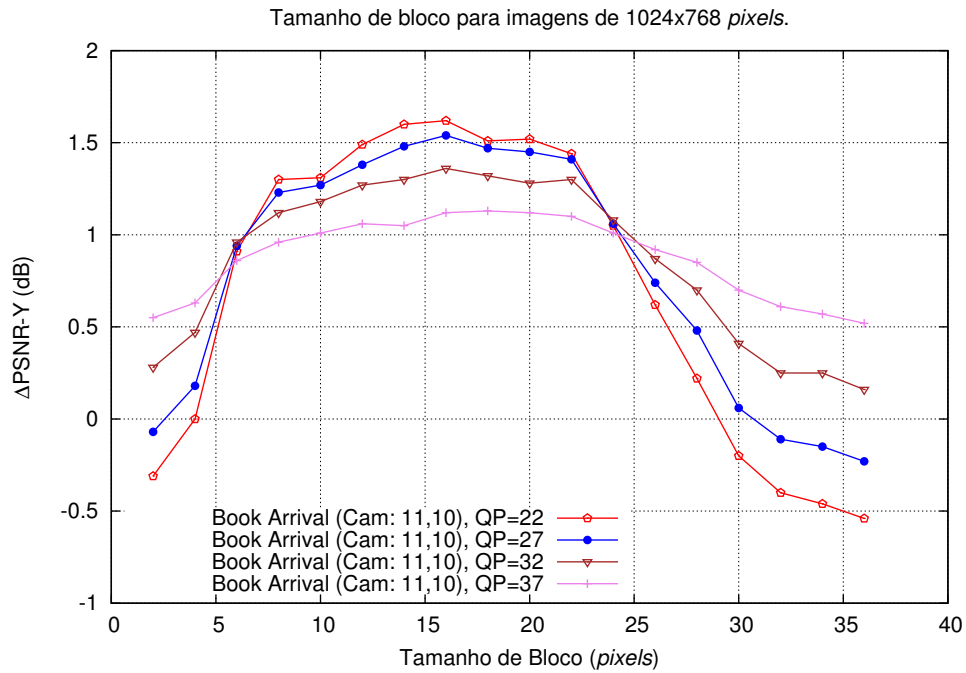


Figura A.38: Diferença de PSNR da sequência *Book Arrival* em função da variação do tamanho do bloco, com a vista auxiliar V.

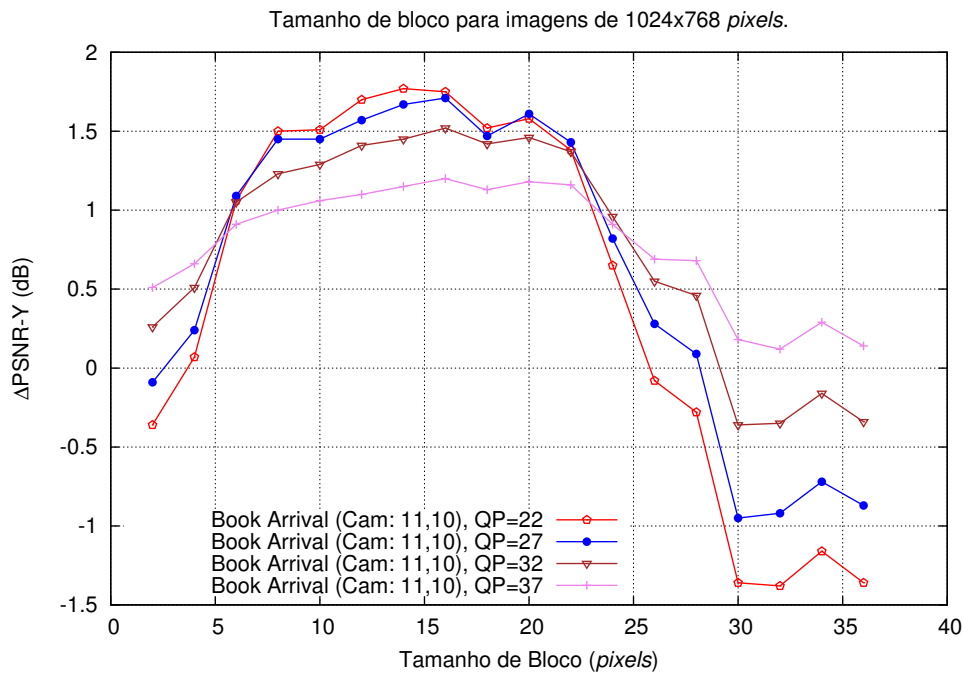


Figura A.39: Diferença de PSNR da sequência *Book Arrival* em função da variação do tamanho do bloco, com a vista auxiliar H/V.

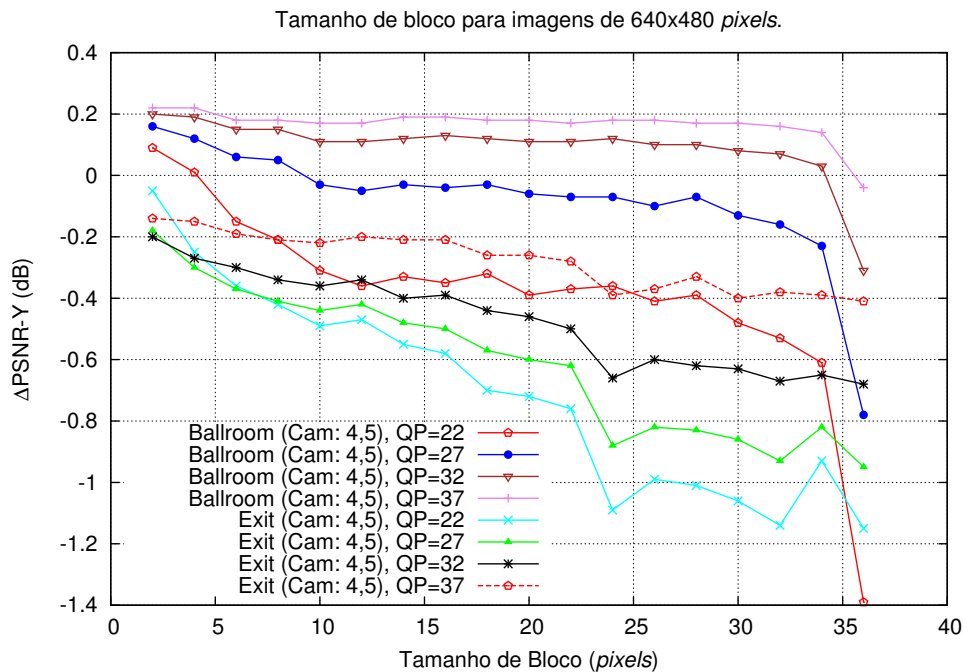


Figura A.40: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho do bloco, com a vista auxiliar H.

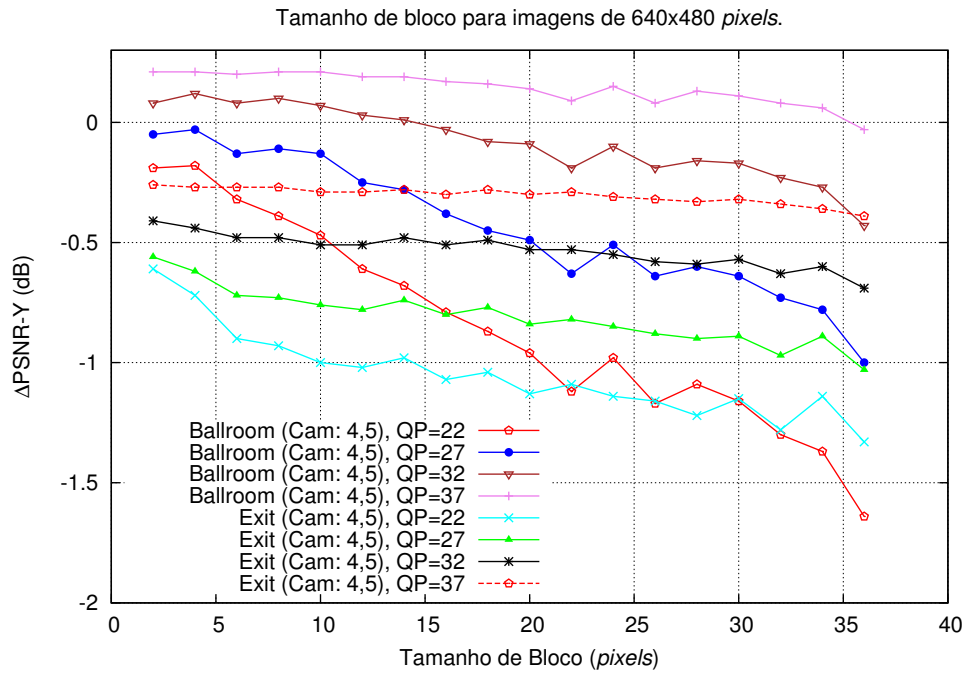


Figura A.41: Diferença de PSNR das seqüências *Ballroom* e *Exit* em função da variação do tamanho do bloco, com a vista auxiliar V.

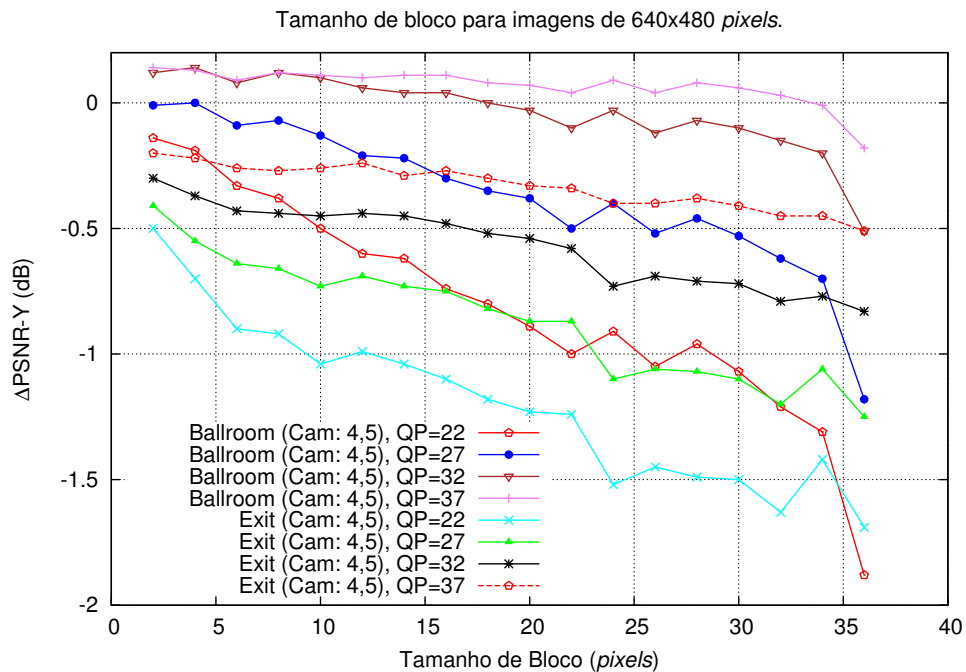


Figura A.42: Diferença de PSNR das seqüências *Ballroom* e *Exit* em função da variação do tamanho do bloco, com a vista auxiliar H/V.

A.4 Largura da janela de pesquisa para o método baseado na pesquisa de blocos

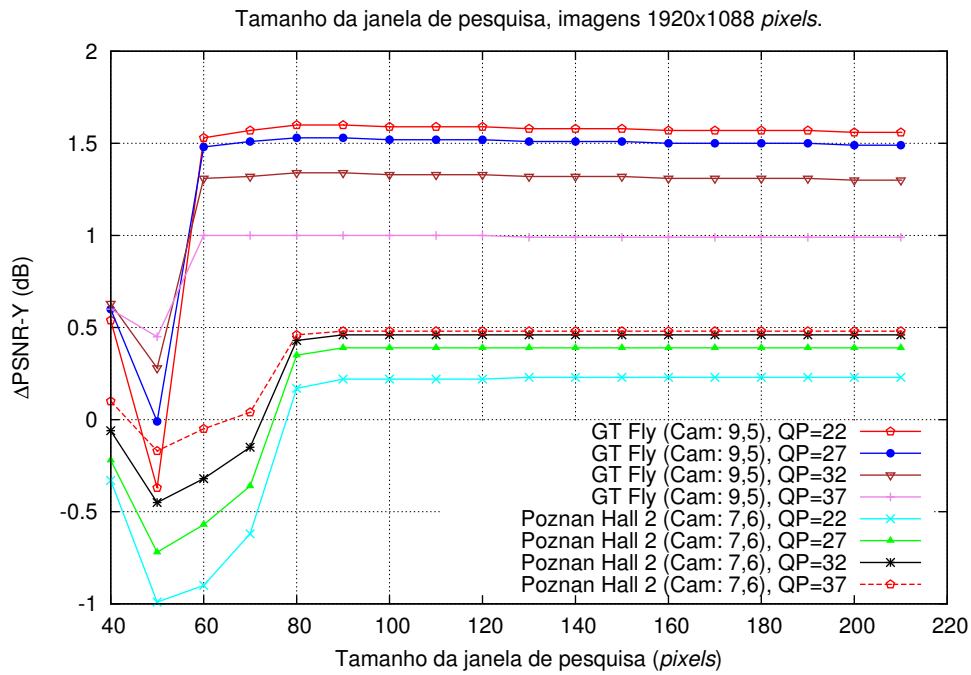


Figura A.43: Diferença de PSNR das sequências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

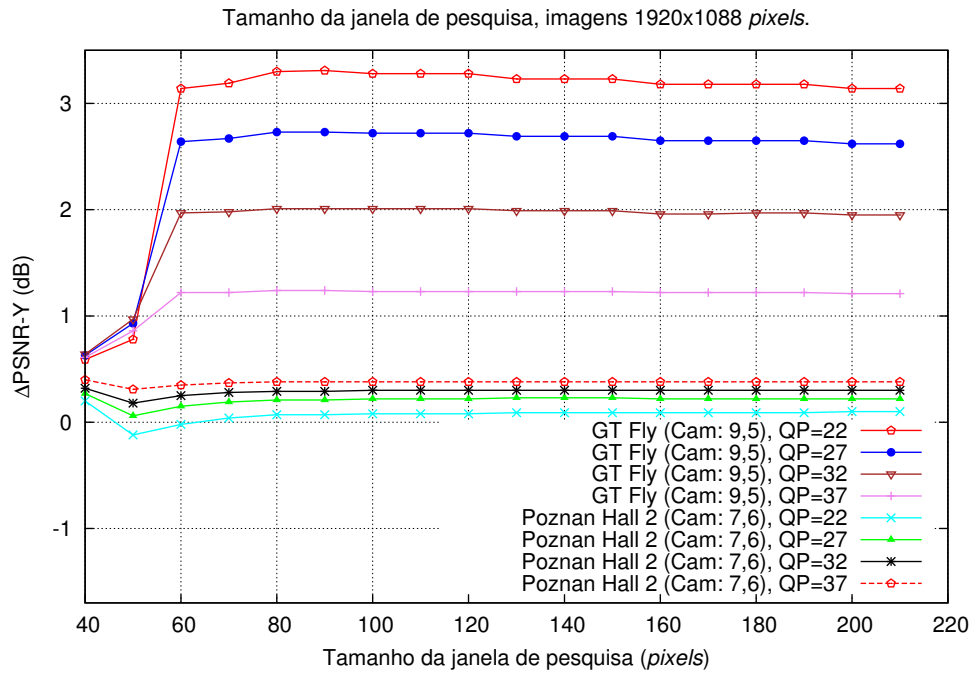


Figura A.44: Diferença de PSNR das seqüências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

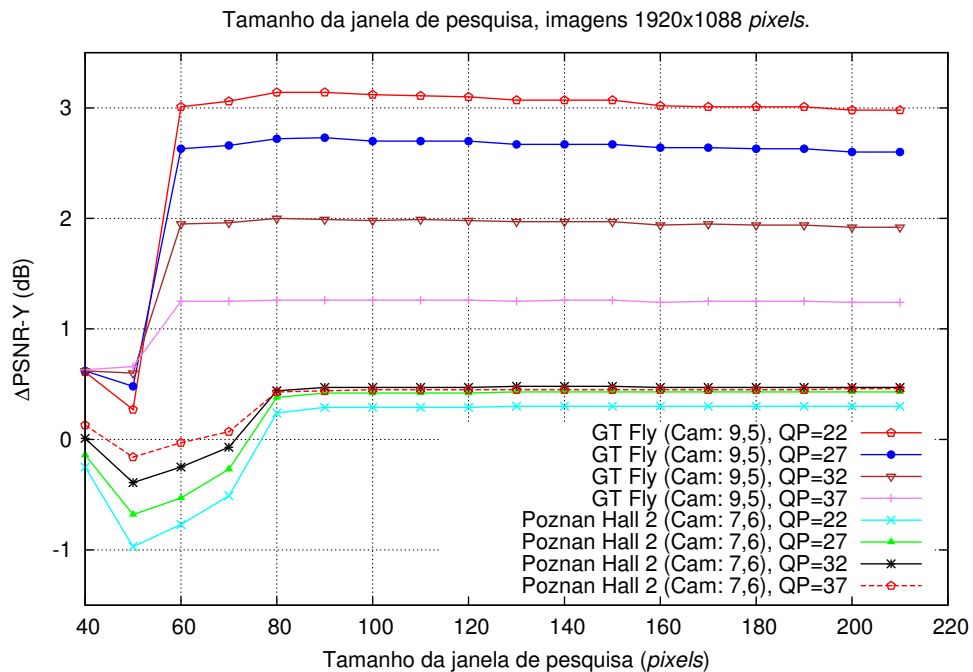


Figura A.45: Diferença de PSNR das seqüências *GT Fly* e *Poznan Hall 2* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

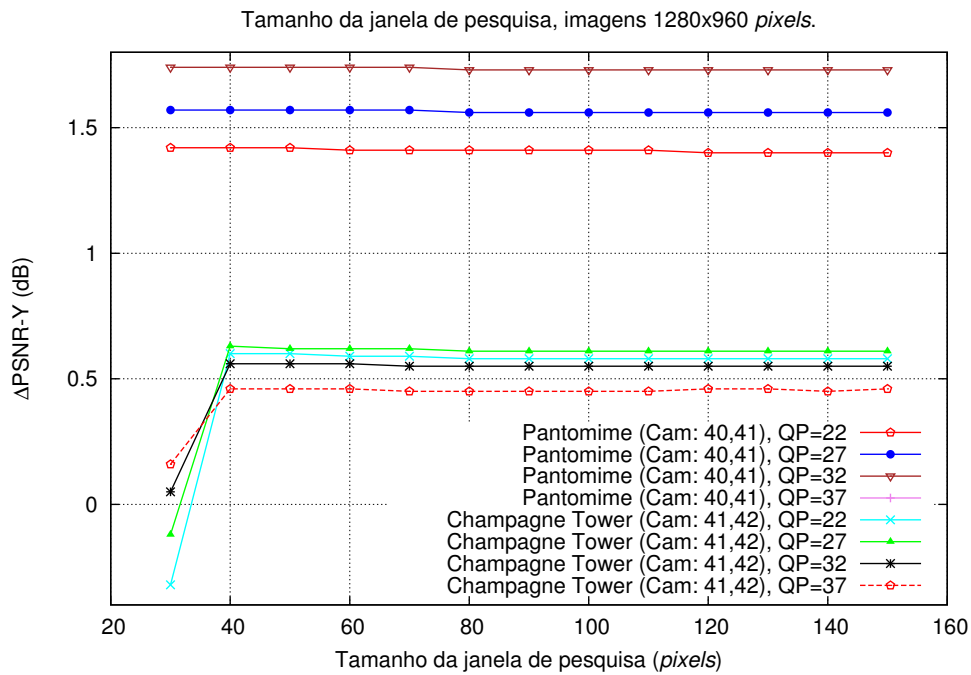


Figura A.46: Diferença de PSNR das seqüências *Pantomime* e *Champagne Tower* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

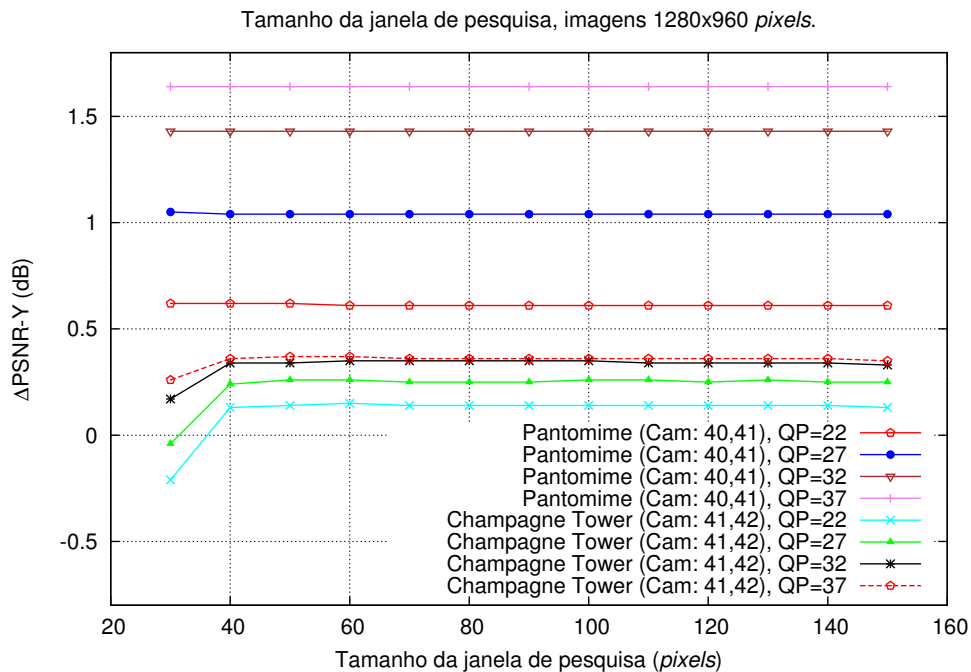


Figura A.47: Diferença de PSNR das seqüências *Pantomime* e *Champagne Tower* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

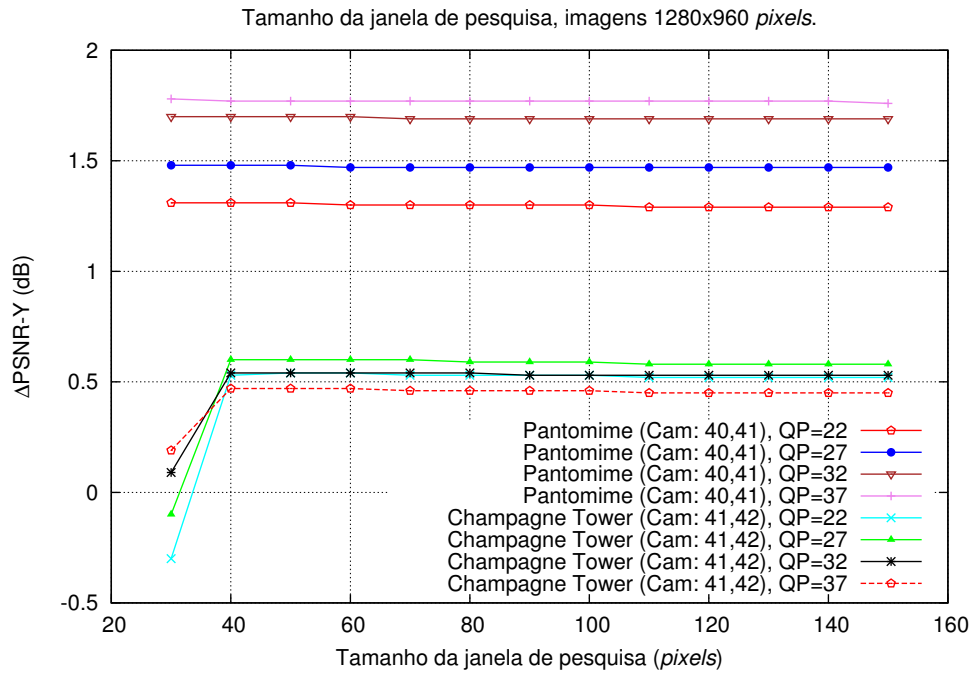


Figura A.48: Diferença de PSNR das sequências *Pantomime* e *Champagne Tower* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

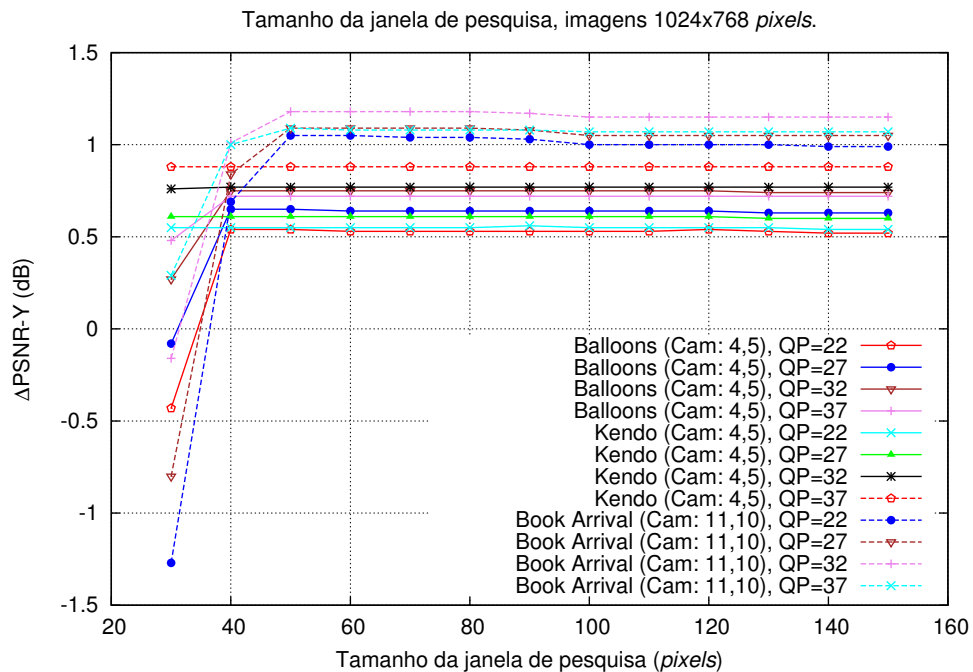


Figura A.49: Diferença de PSNR das sequências *Balloons*, *Kendo* e *Book Arrival* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

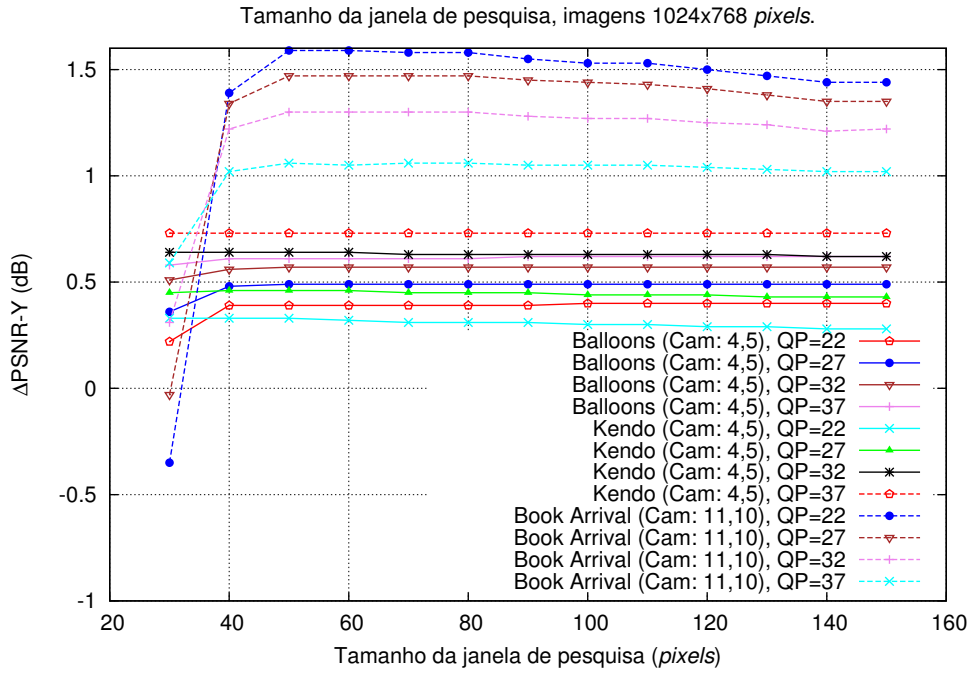


Figura A.50: Diferença de PSNR das sequências *Balloons*, *Kendo* e *Book Arrival* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

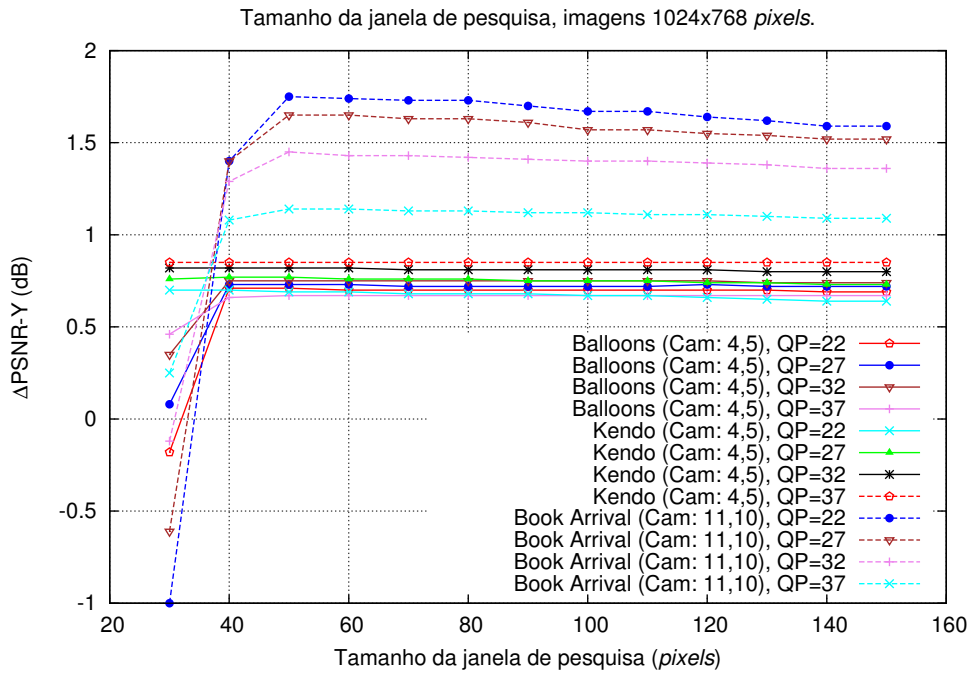


Figura A.51: Diferença de PSNR das sequências *Balloons*, *Kendo* e *Book Arrival* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

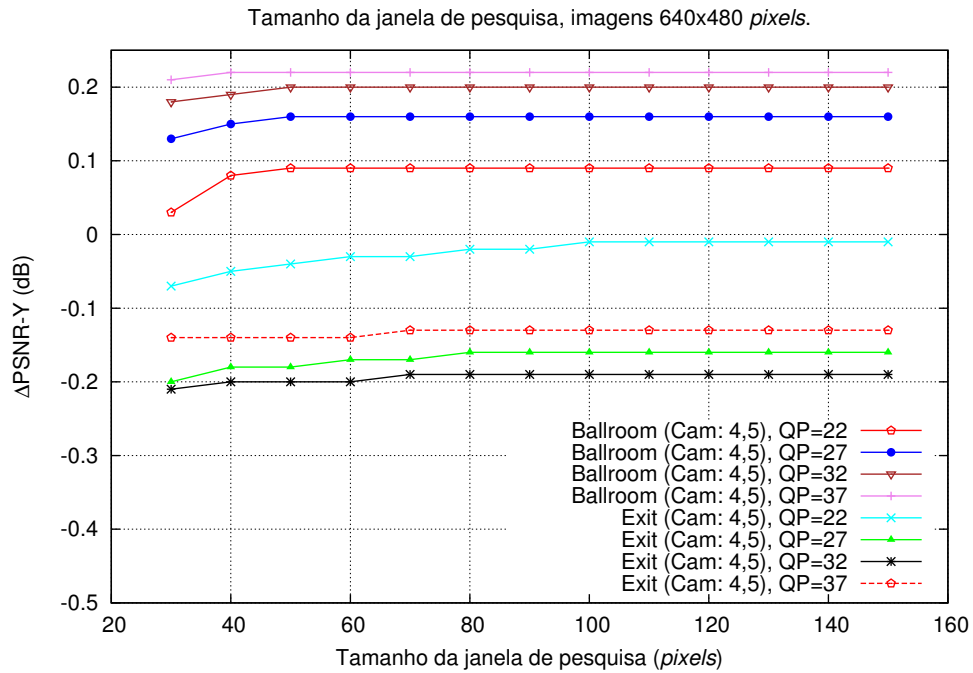


Figura A.52: Diferença de PSNR das seqüências *Ballroom* e *Exit* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H.

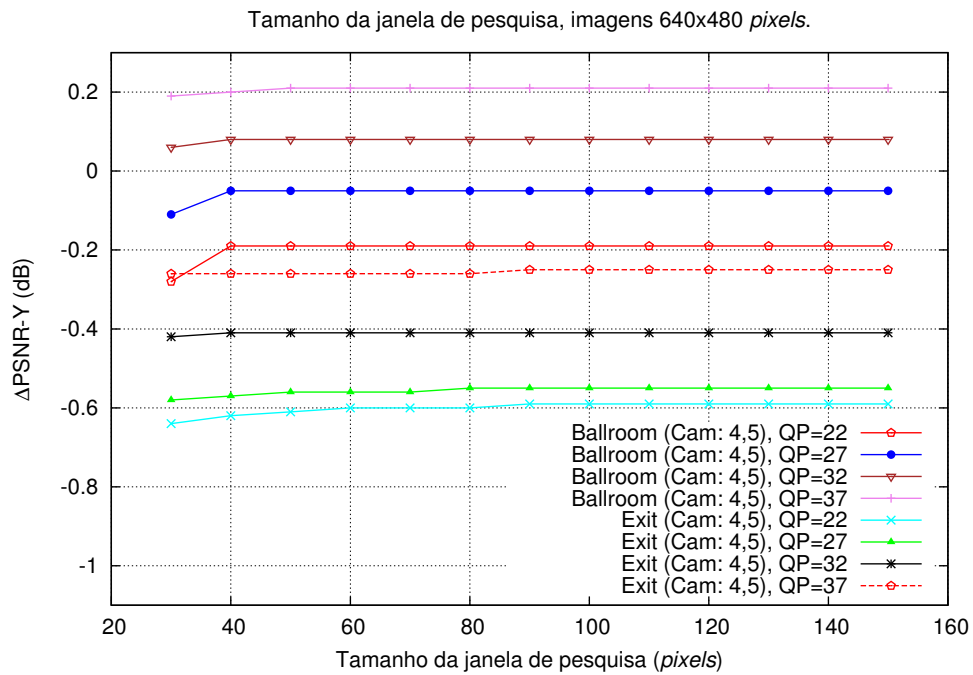


Figura A.53: Diferença de PSNR das seqüências *Ballroom* e *Exit* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar V.

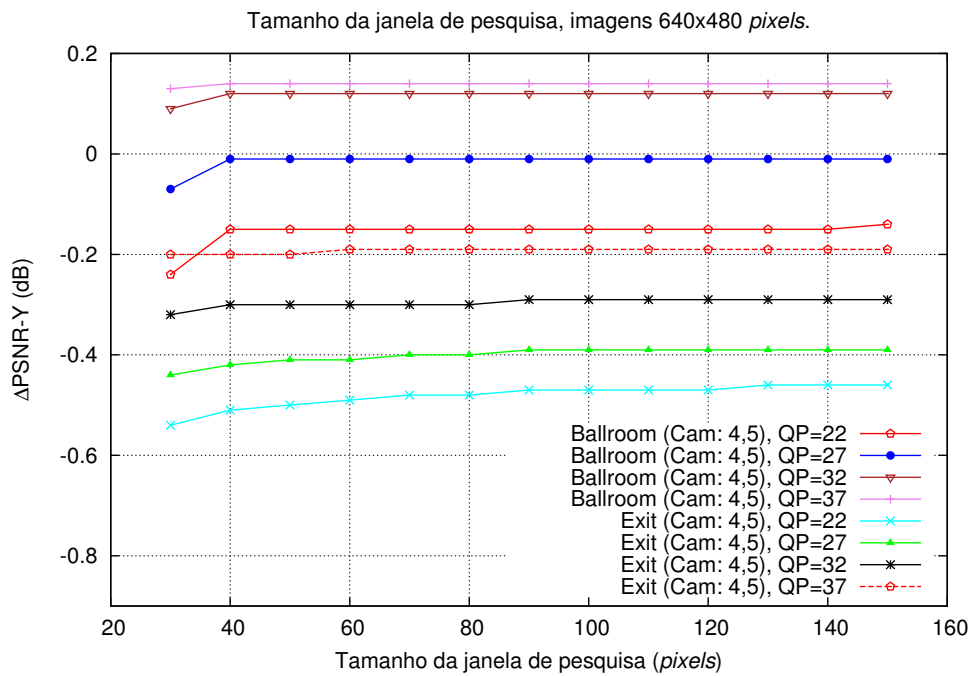


Figura A.54: Diferença de PSNR das sequências *Ballroom* e *Exit* em função da variação do tamanho da janela de pesquisa, com a vista auxiliar H/V.

A.5 Vista de referência na pesquisa de blocos

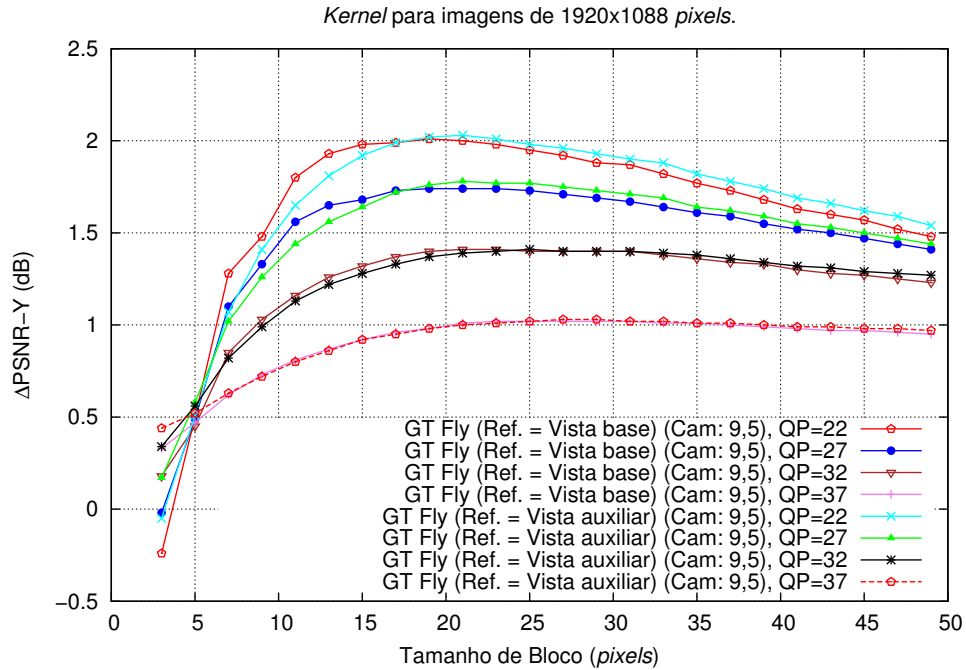


Figura A.55: Seleção da vista de referência, no método da malha, para a sequência *GT Fly*, com a vista auxiliar H.

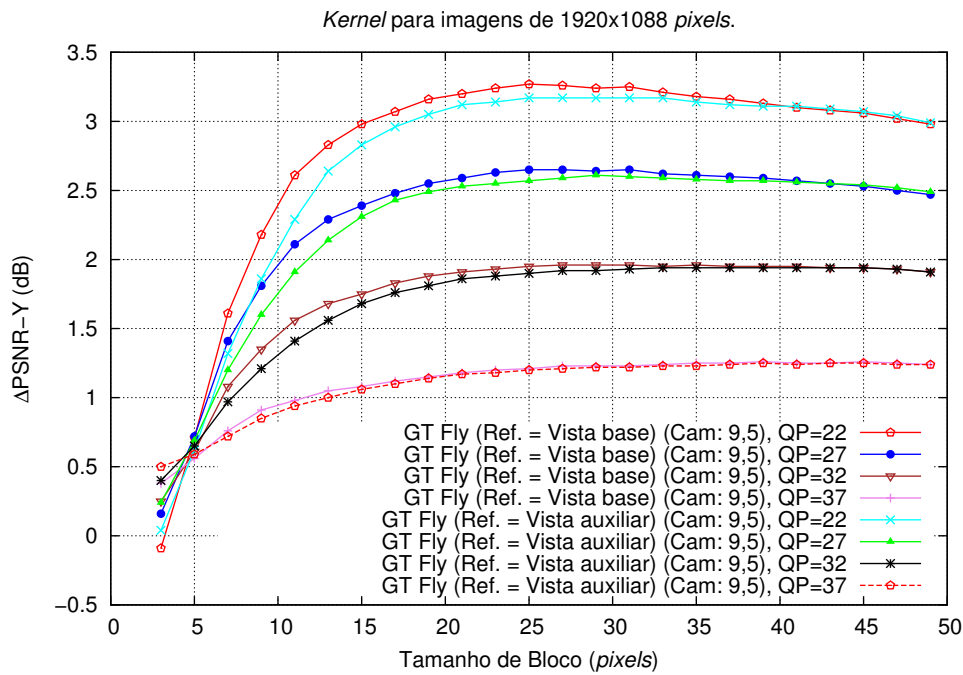


Figura A.56: Seleção da vista de referência, no método da malha, para a sequência *GT Fly*, com a vista auxiliar V.

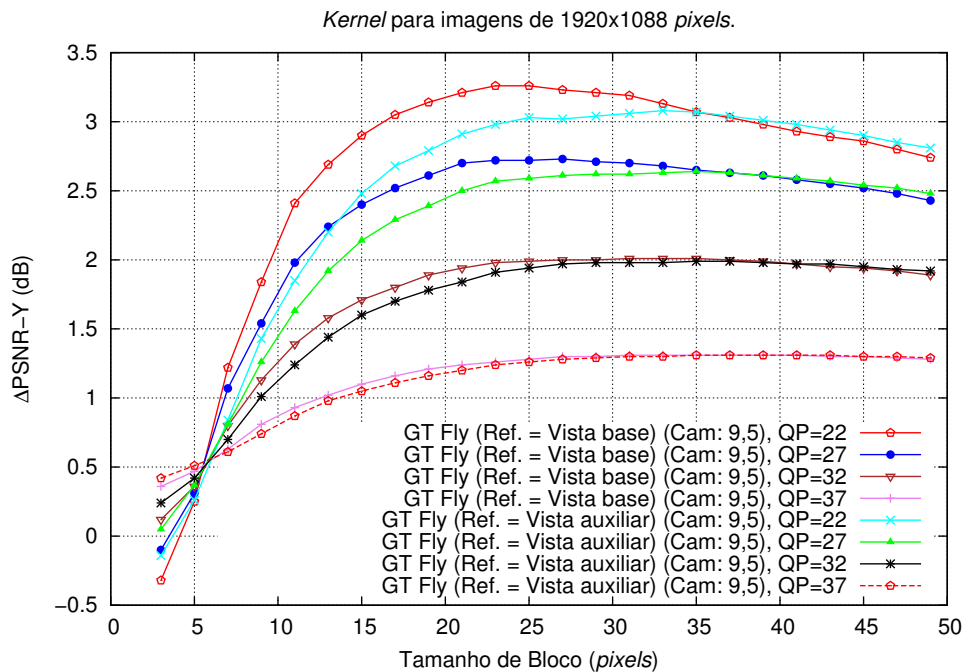


Figura A.57: Seleção da vista de referência, no método da malha, para a sequência *GT Fly*, com a vista auxiliar H/V.

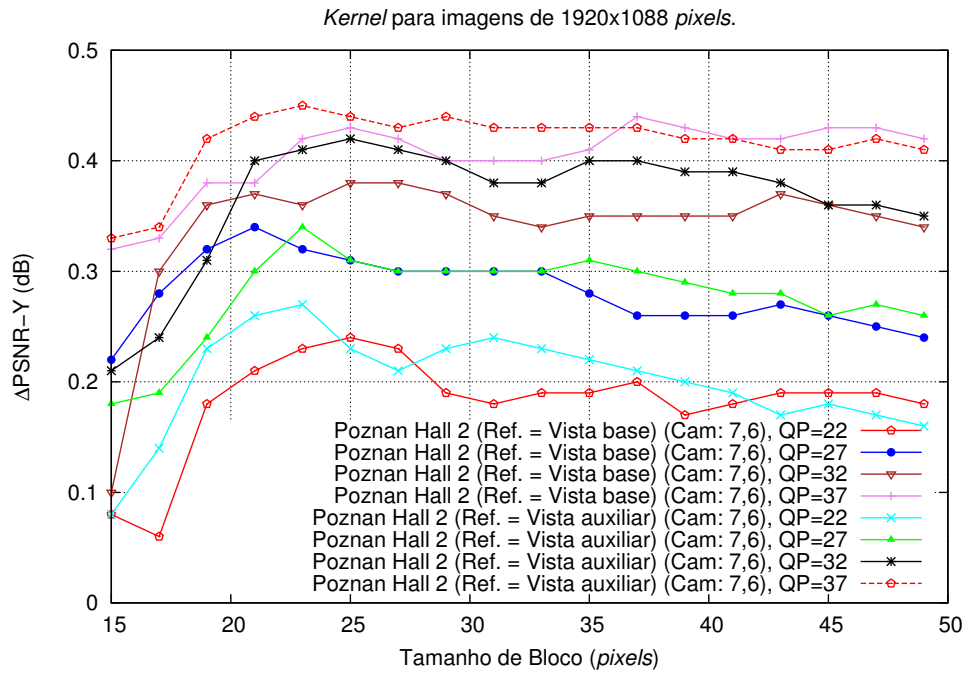


Figura A.58: Seleção da vista de referência, no método da malha, para a sequência *Poznan Hall 2*, com a vista auxiliar H.

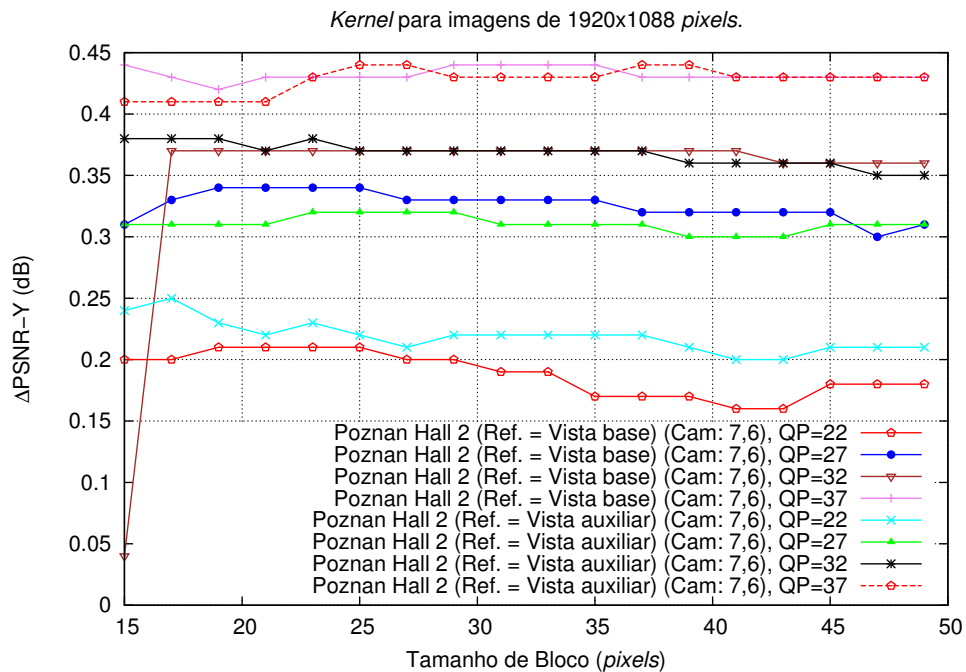


Figura A.59: Seleção da vista de referência, no método da malha, para a sequência *Poznan Hall 2*, com a vista auxiliar V.

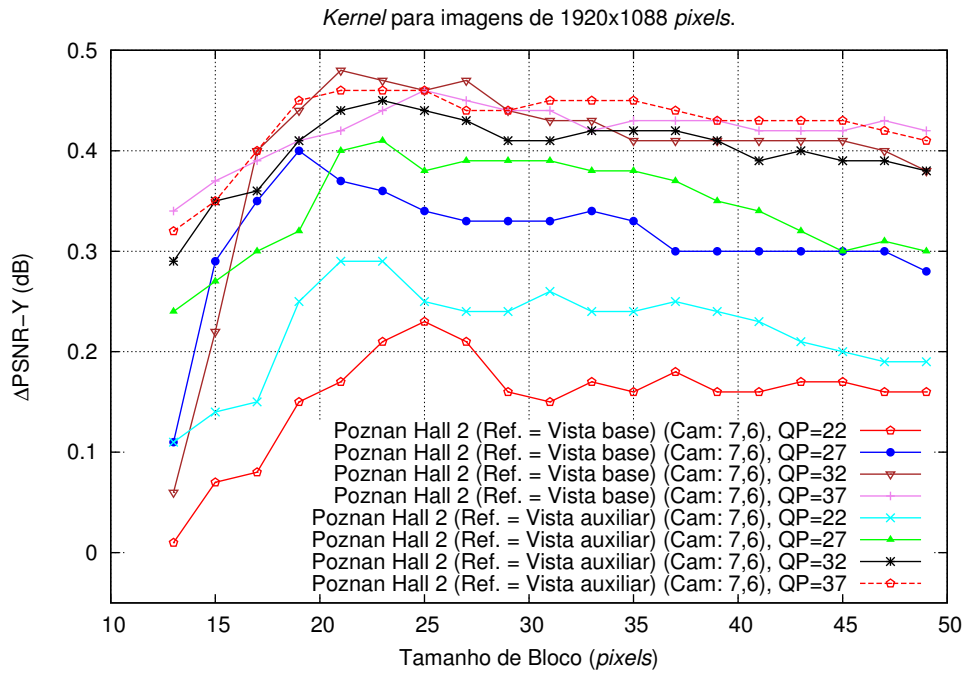


Figura A.60: Seleção da vista de referência, no método da malha, para a sequência *Poznan Hall 2*, com a vista auxiliar H/V.

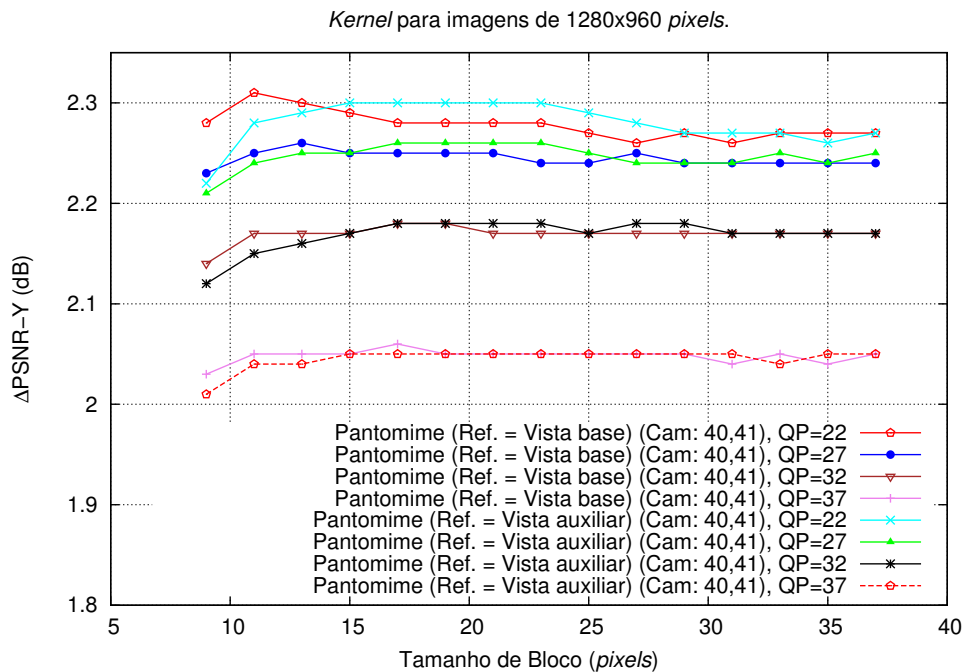


Figura A.61: Seleção da vista de referência, no método da malha, para a sequência *Pantomime*, com a vista auxiliar H.

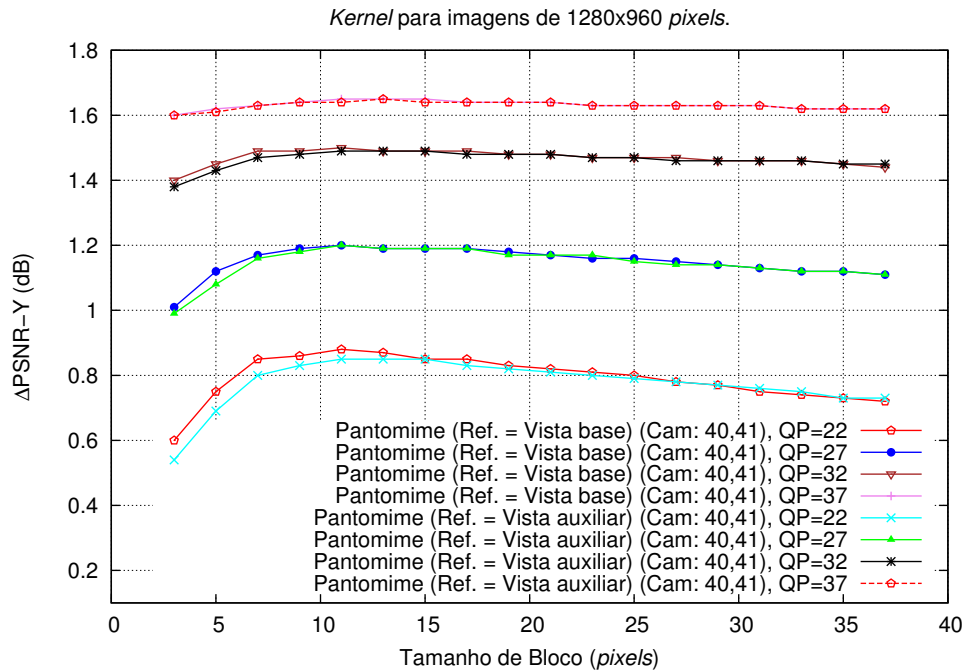


Figura A.62: Seleção da vista de referência, no método da malha, para a seqüência *Pantomime*, com a vista auxiliar V.

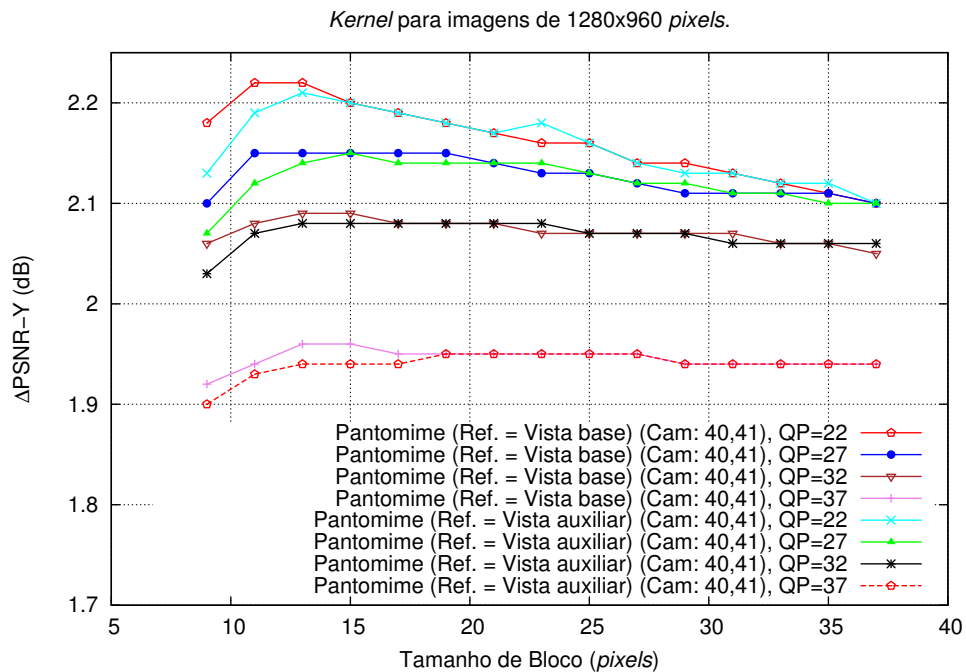


Figura A.63: Seleção da vista de referência, no método da malha, para a seqüência *Pantomime*, com a vista auxiliar H/V.

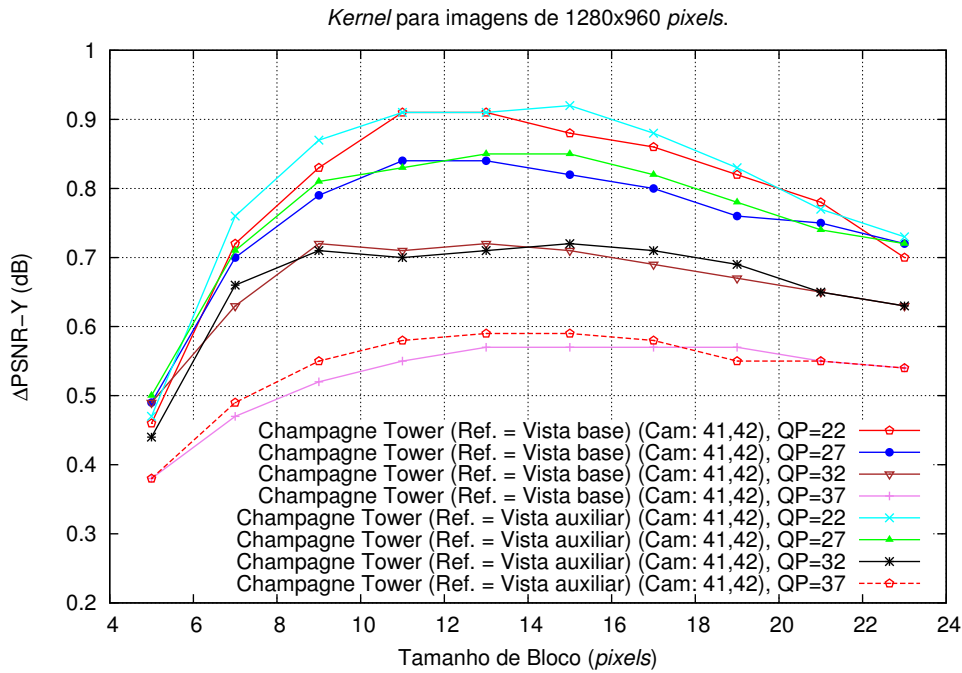


Figura A.64: Seleção da vista de referência, no método da malha, para a sequência *Champagne Tower*, com a vista auxiliar H.

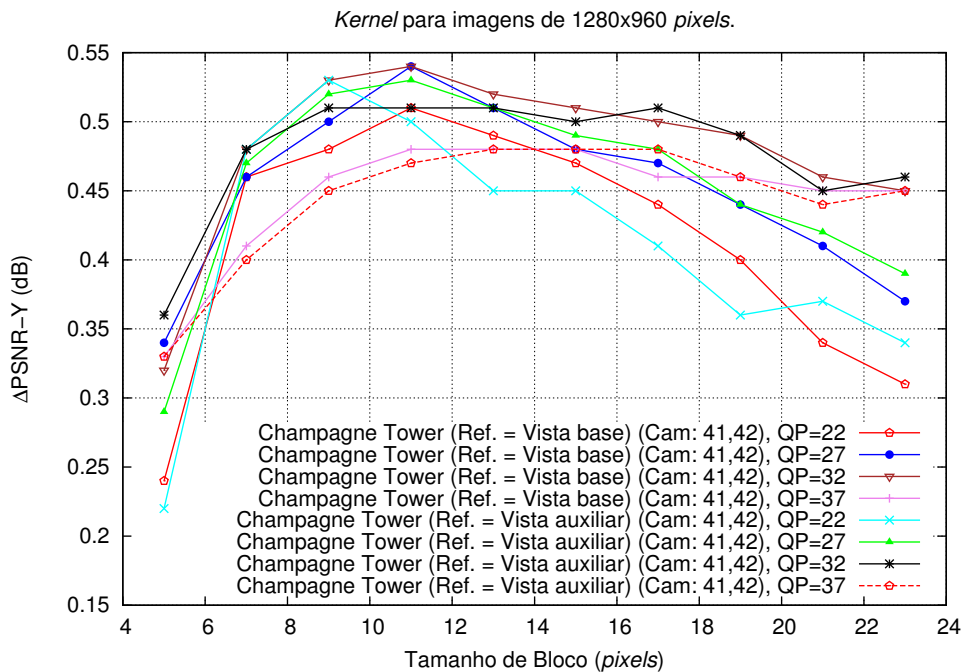


Figura A.65: Seleção da vista de referência, no método da malha, para a sequência *Champagne Tower*, com a vista auxiliar V.

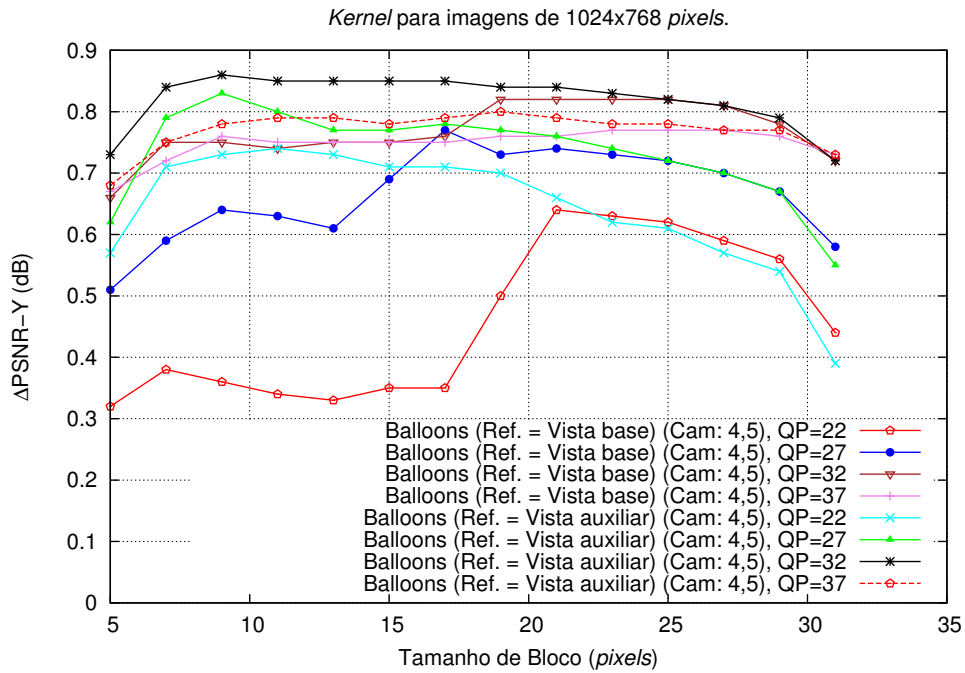


Figura A.66: Seleção da vista de referência, no método da malha, para a sequência *Balloons*, com a vista auxiliar H.

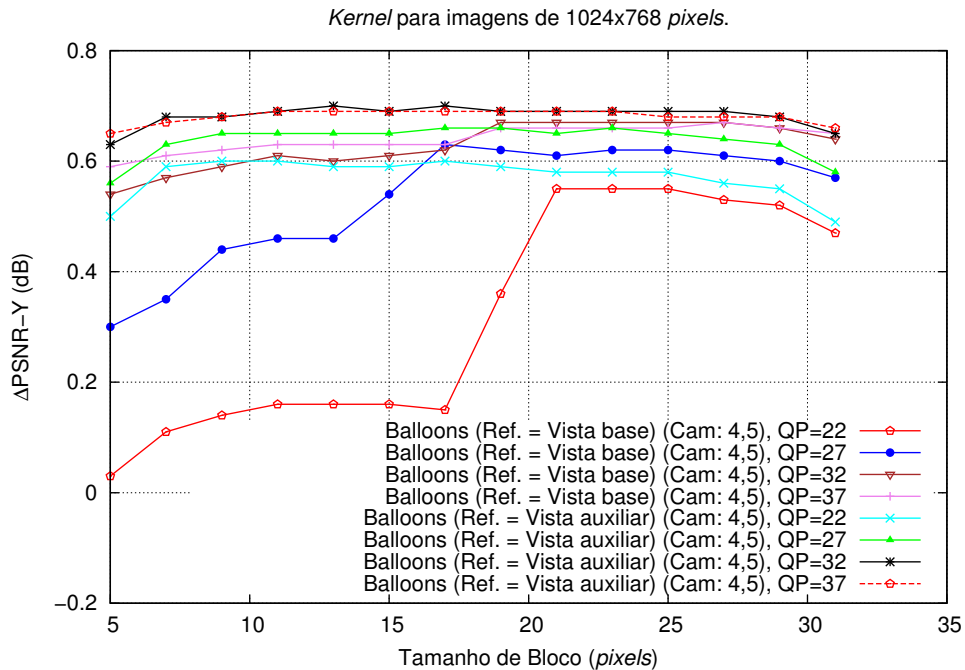


Figura A.67: Seleção da vista de referência, no método da malha, para a sequência *Balloons*, com a vista auxiliar V.

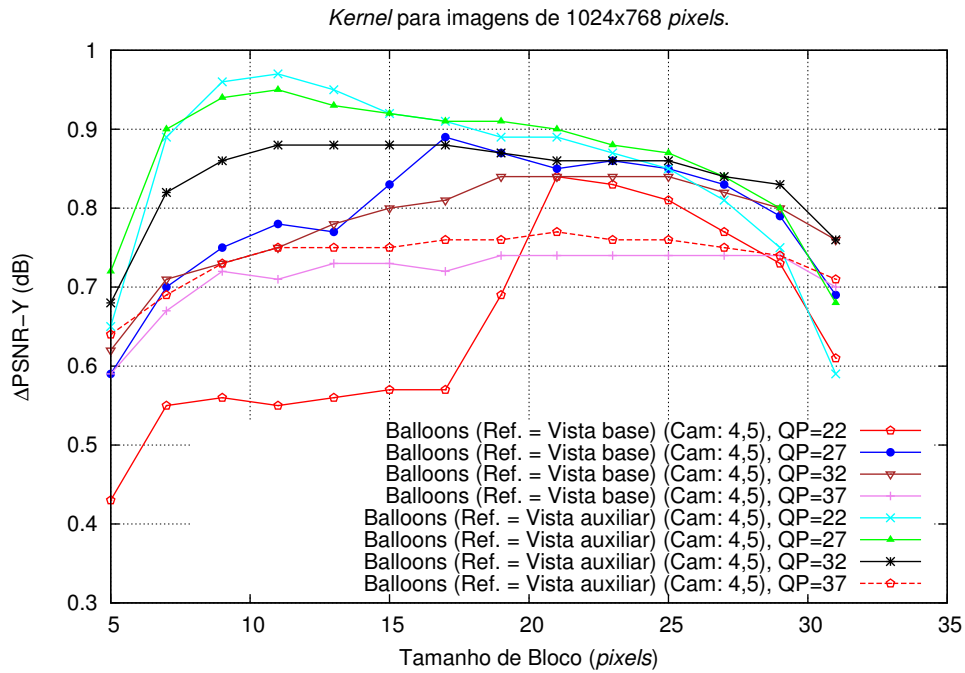


Figura A.68: Seleção da vista de referência, no método da malha, para a sequência *Balloons*, com a vista auxiliar H/V.

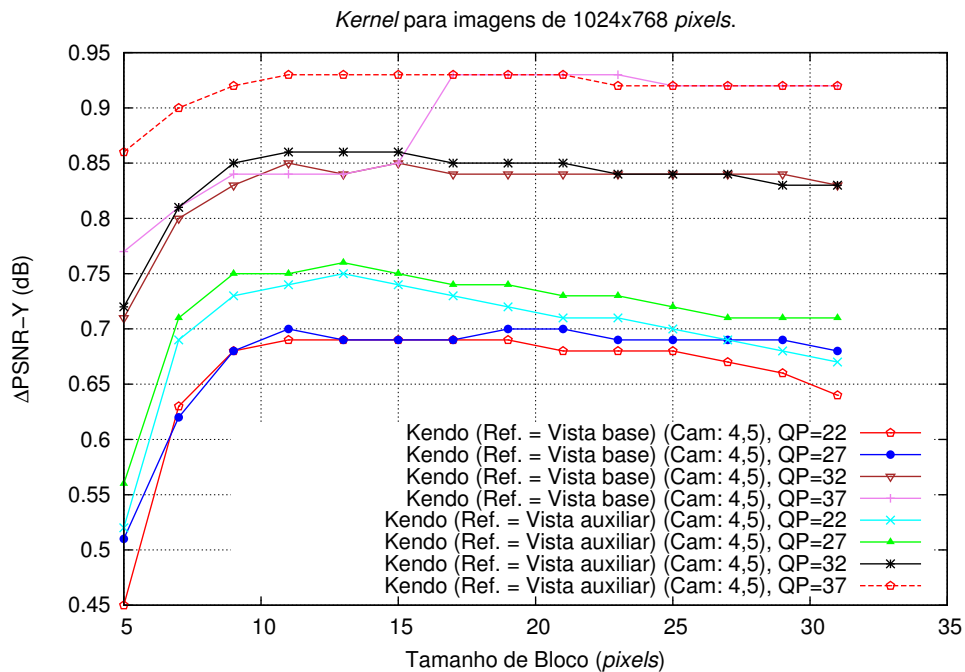


Figura A.69: Seleção da vista de referência, no método da malha, para a sequência *Kendo*, com a vista auxiliar H.

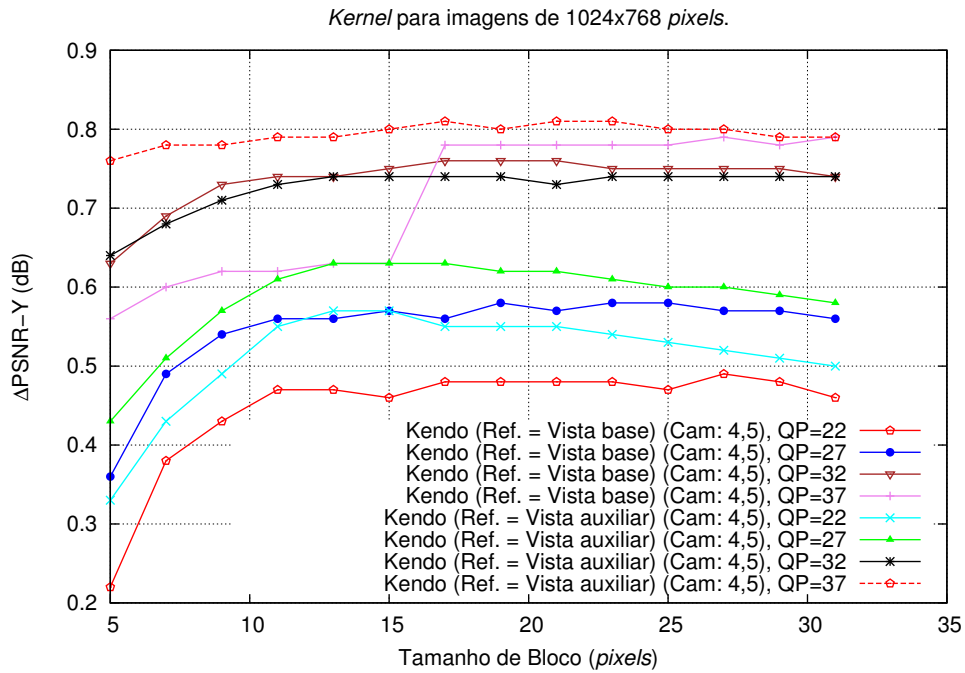


Figura A.70: Seleção da vista de referência, no método da malha, para a sequência *Kendo*, com a vista auxiliar V.

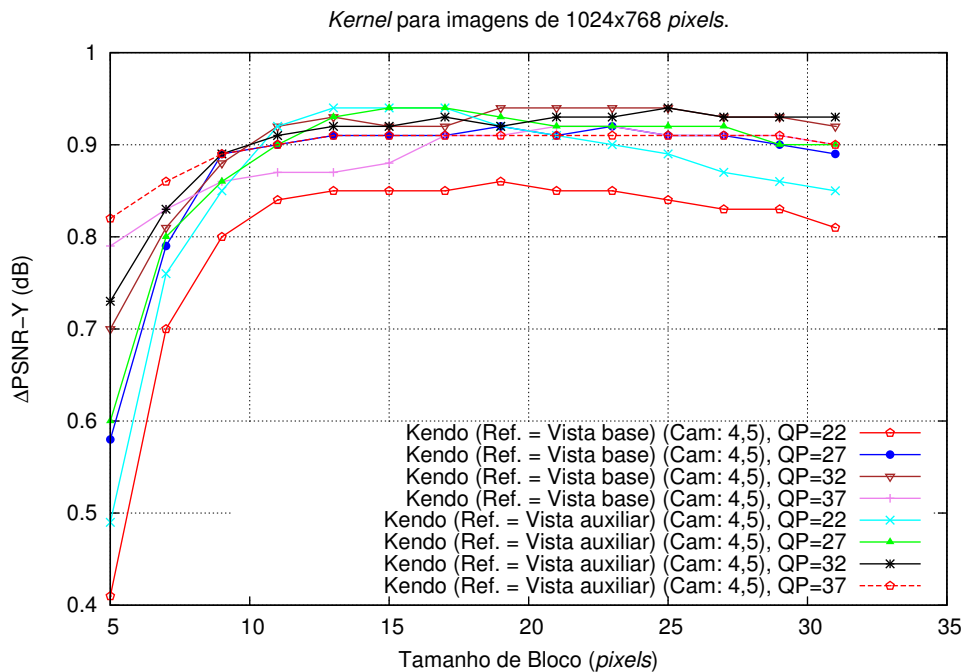


Figura A.71: Seleção da vista de referência, no método da malha, para a sequência *Kendo*, com a vista auxiliar H/V.

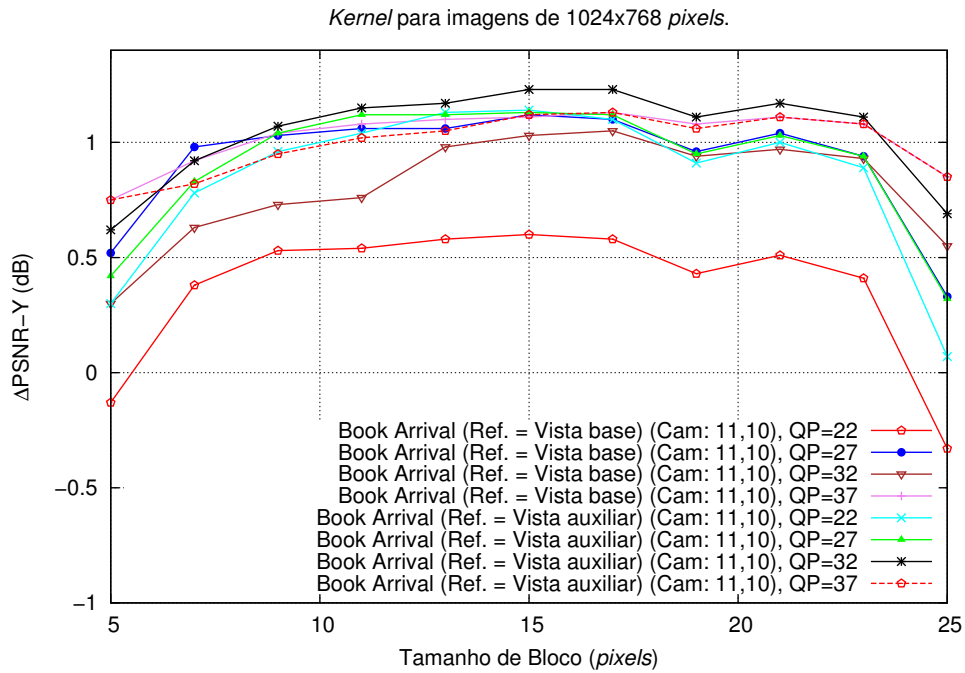


Figura A.72: Seleção da vista de referência, no método da malha, para a sequência *Book Arrival*, com a vista auxiliar H.

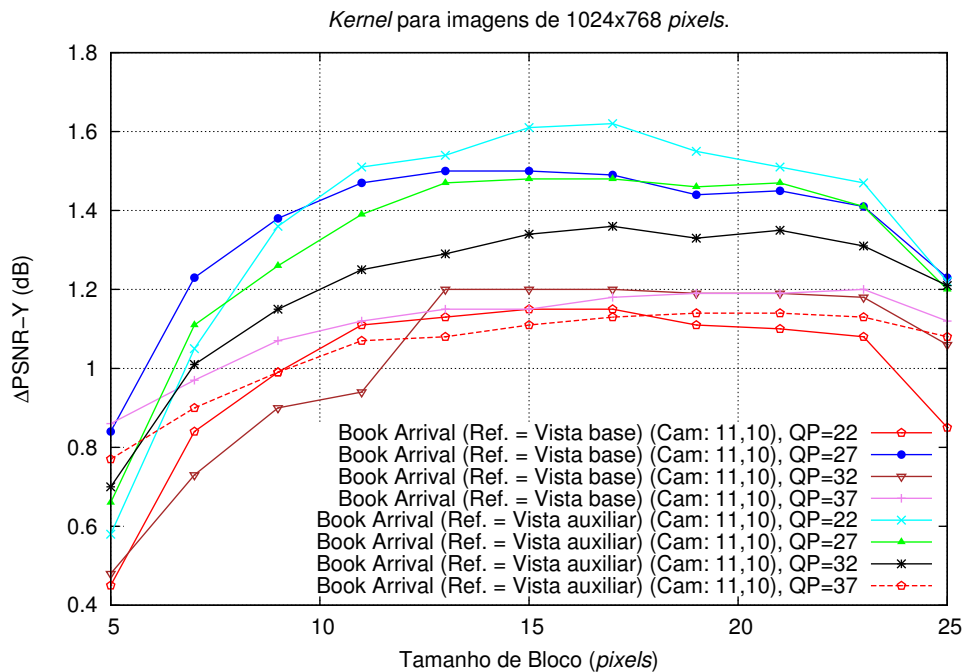


Figura A.73: Seleção da vista de referência, no método da malha, para a sequência *Book Arrival*, com a vista auxiliar V.

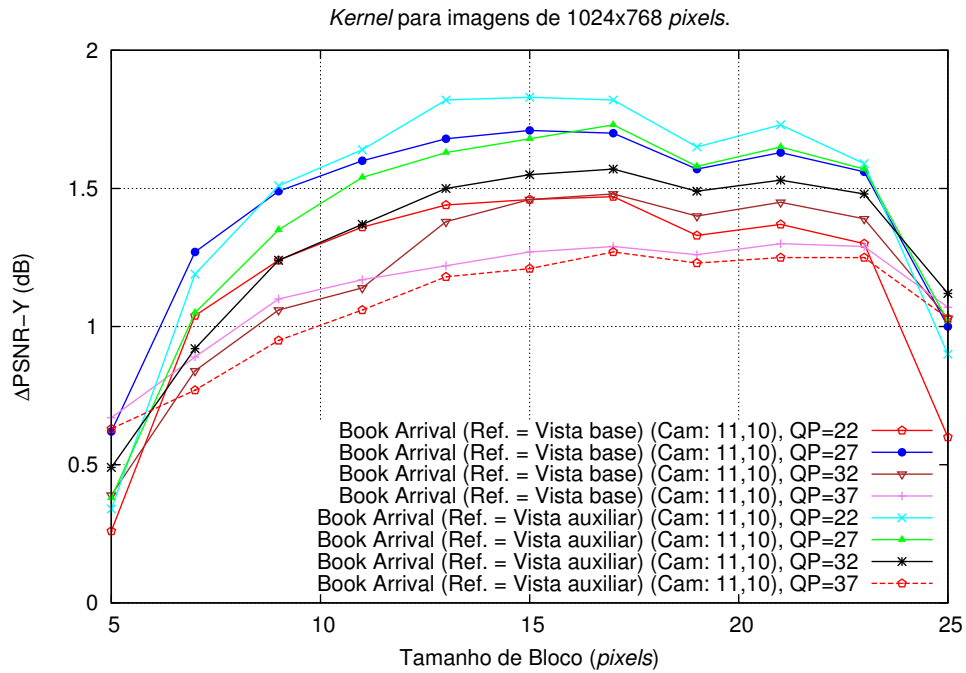


Figura A.74: Seleção da vista de referência, no método da malha, para a sequência *Book Arrival*, com a vista auxiliar H/V.

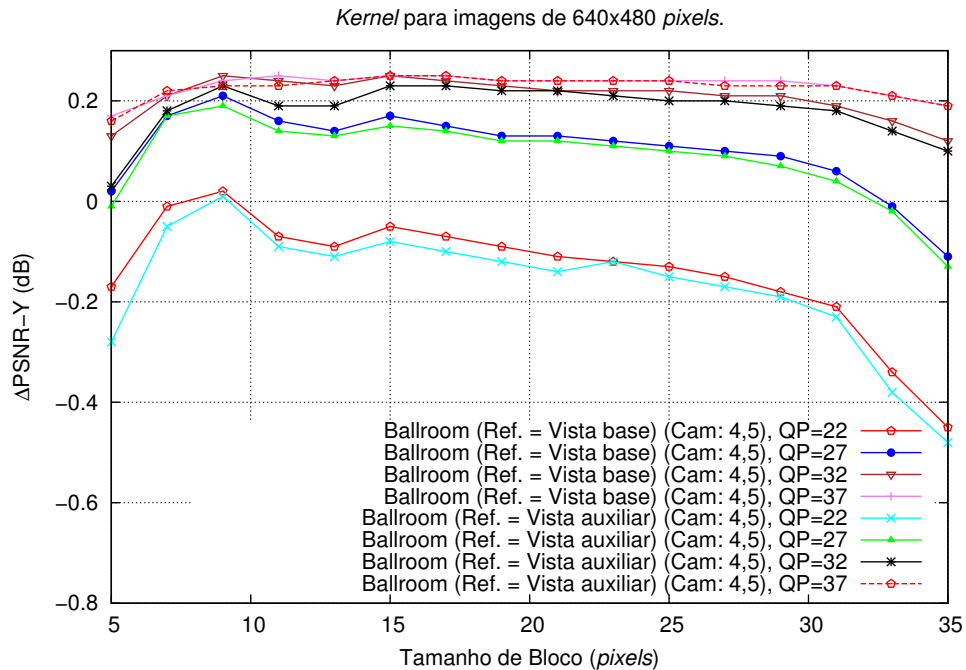


Figura A.75: Seleção da vista de referência, no método da malha, para a sequência *Ballroom*, com a vista auxiliar H.

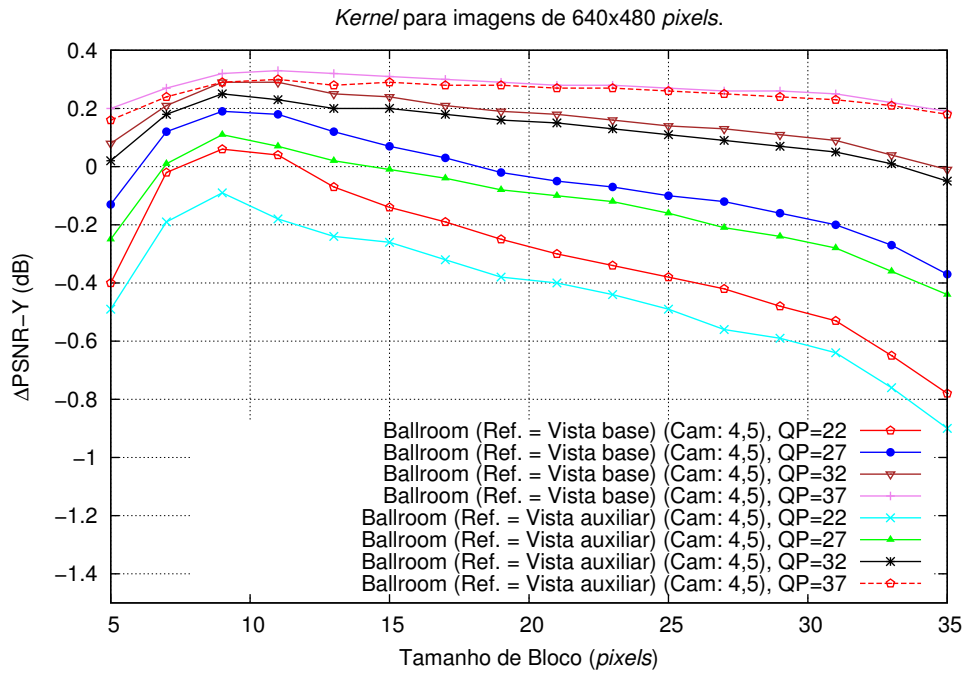


Figura A.76: Seleção da vista de referência, no método da malha, para a sequência *Ballroom*, com a vista auxiliar V.

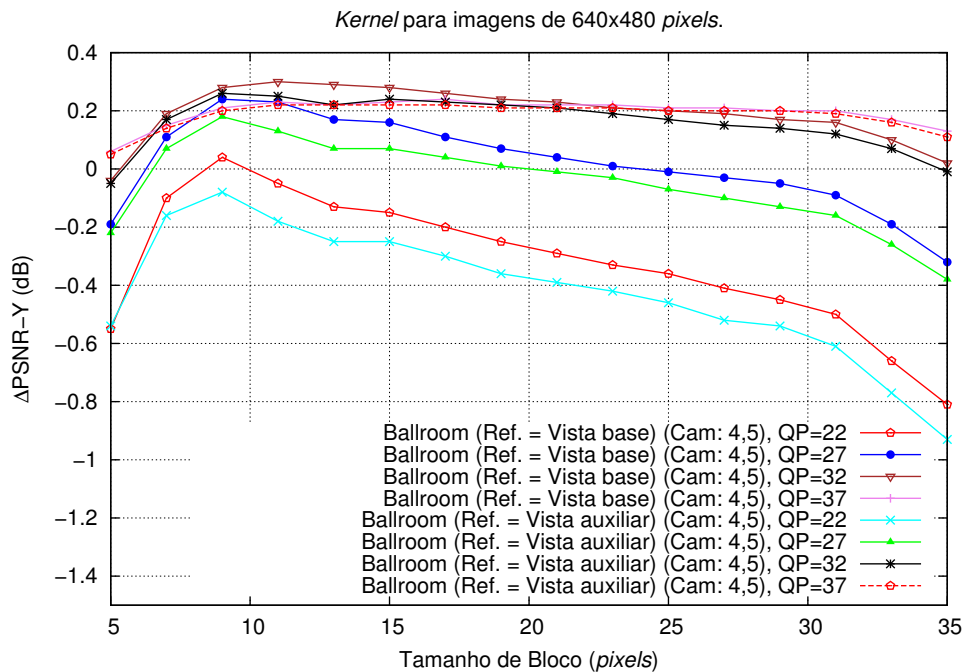


Figura A.77: Seleção da vista de referência, no método da malha, para a sequência *Ballroom*, com a vista auxiliar H/V.

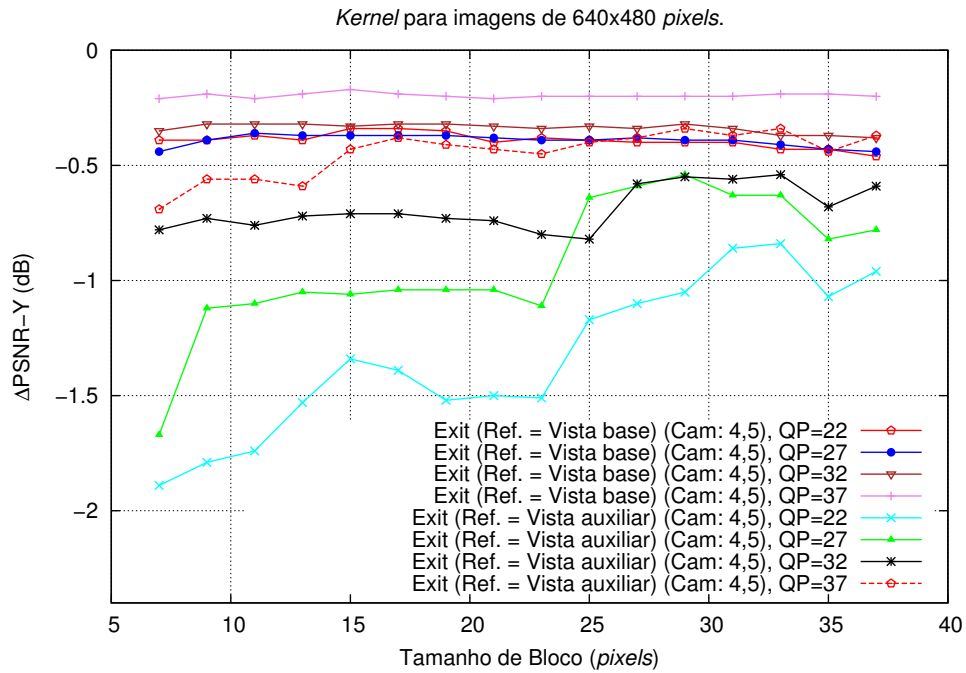


Figura A.78: Seleção da vista de referência, no método da malha, para a sequência *Exit*, com a vista auxiliar H.

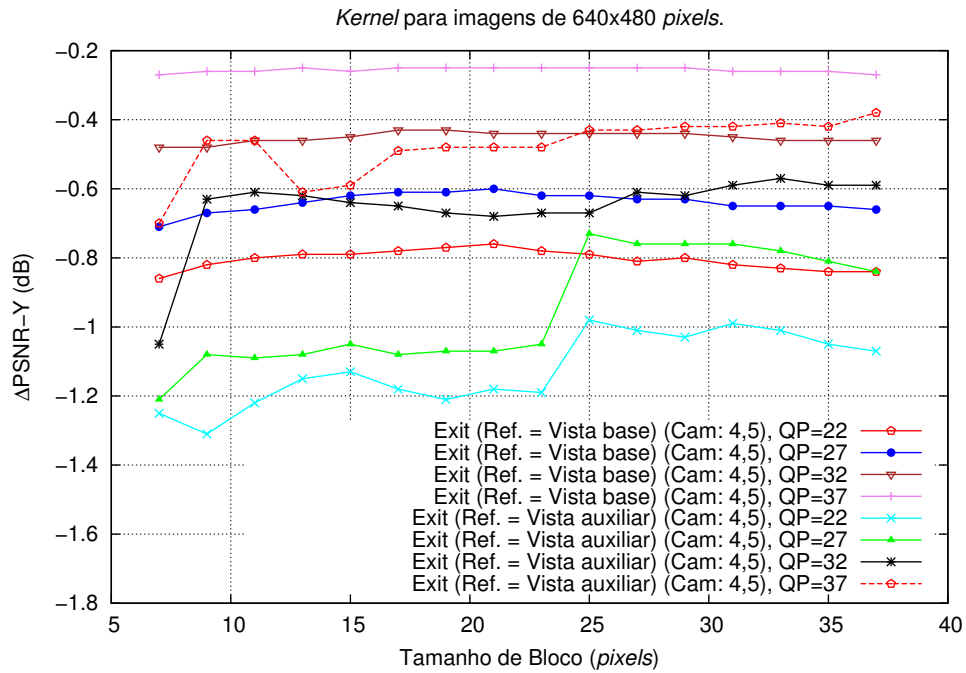


Figura A.79: Seleção da vista de referência, no método da malha, para a sequência *Exit*, com a vista auxiliar V.

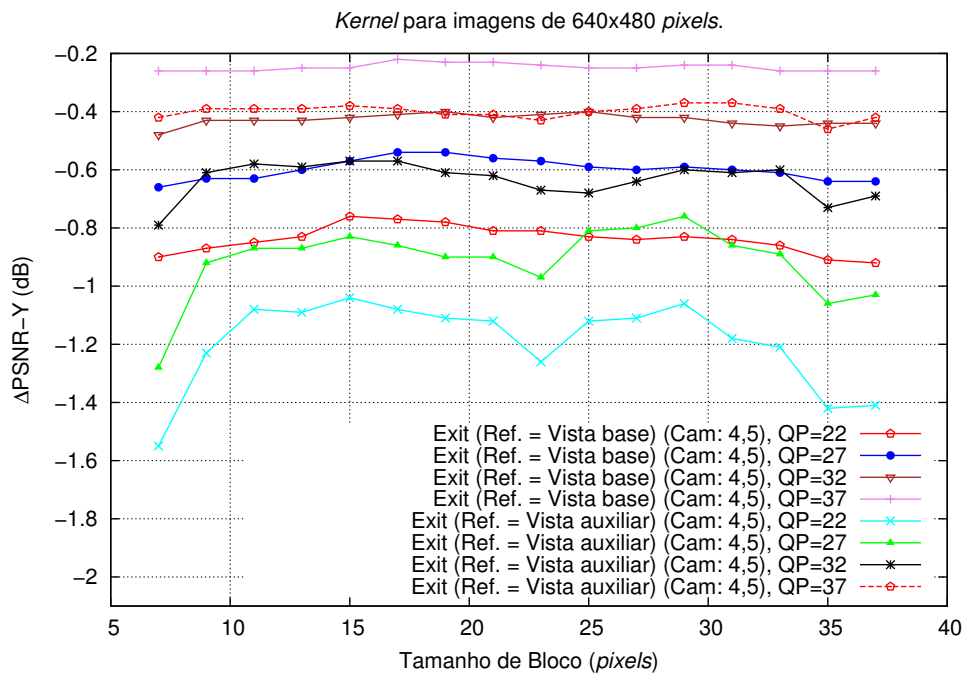


Figura A.80: Seleção da vista de referência, no método da malha, para a sequência *Exit*, com a vista auxiliar H/V.

A.6 Avaliação dos métodos de adição de detalhes

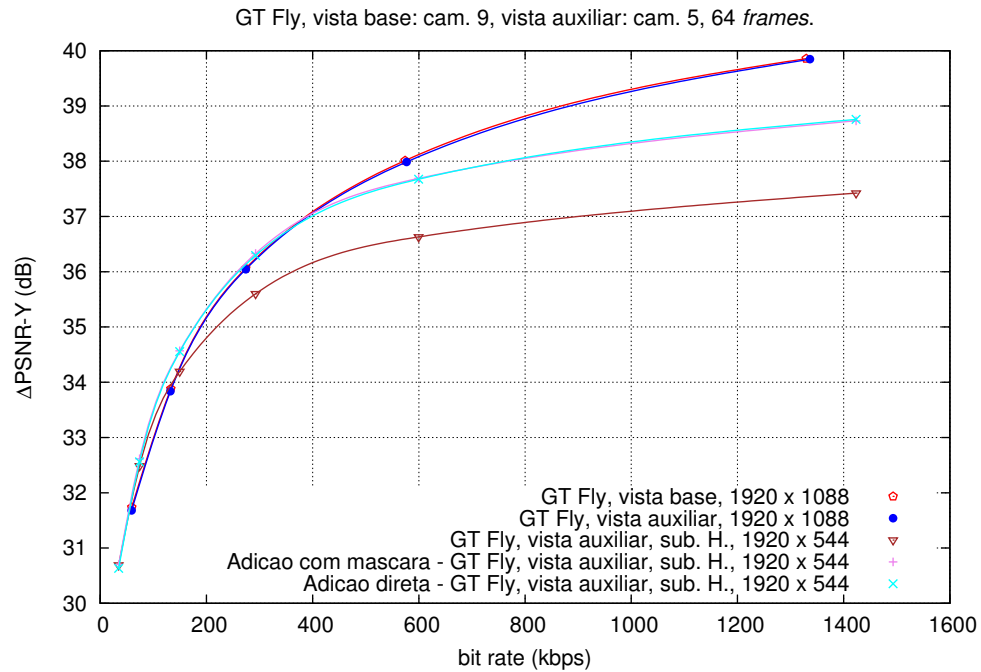


Figura A.81: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *GT Fly*.

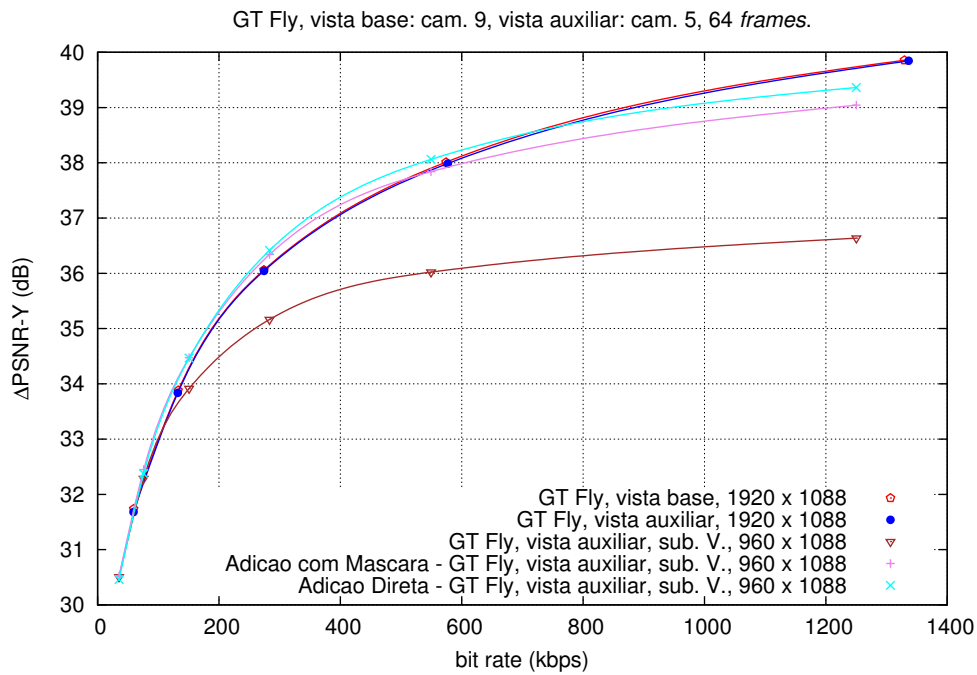


Figura A.82: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V , para a sequência *GT Fly*.

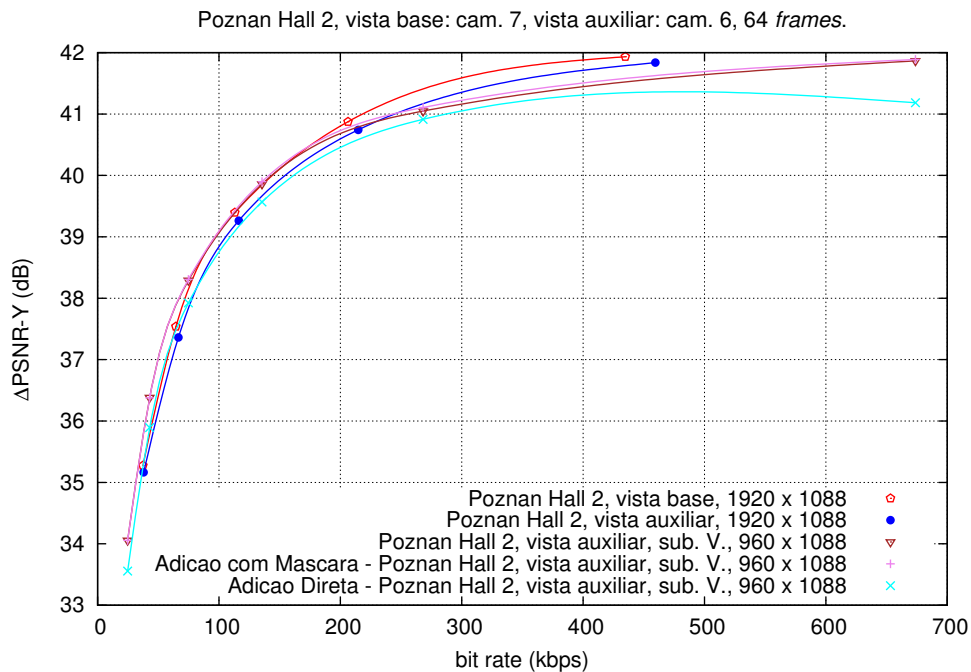


Figura A.83: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V , para a sequência *Poznan Hall 2*.

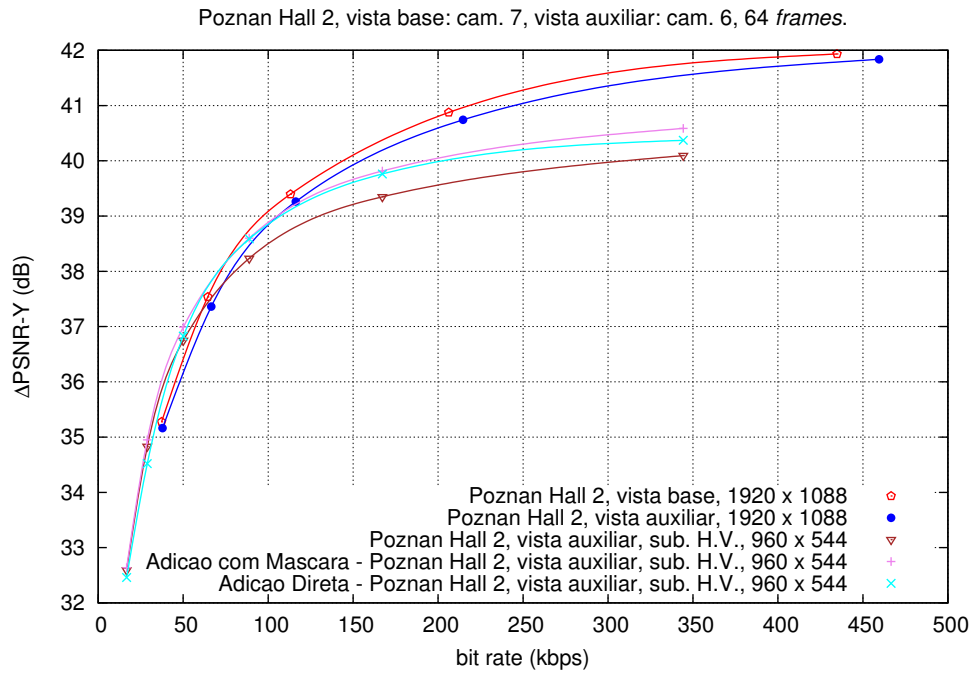


Figura A.84: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência *Poznan Hall 2*.

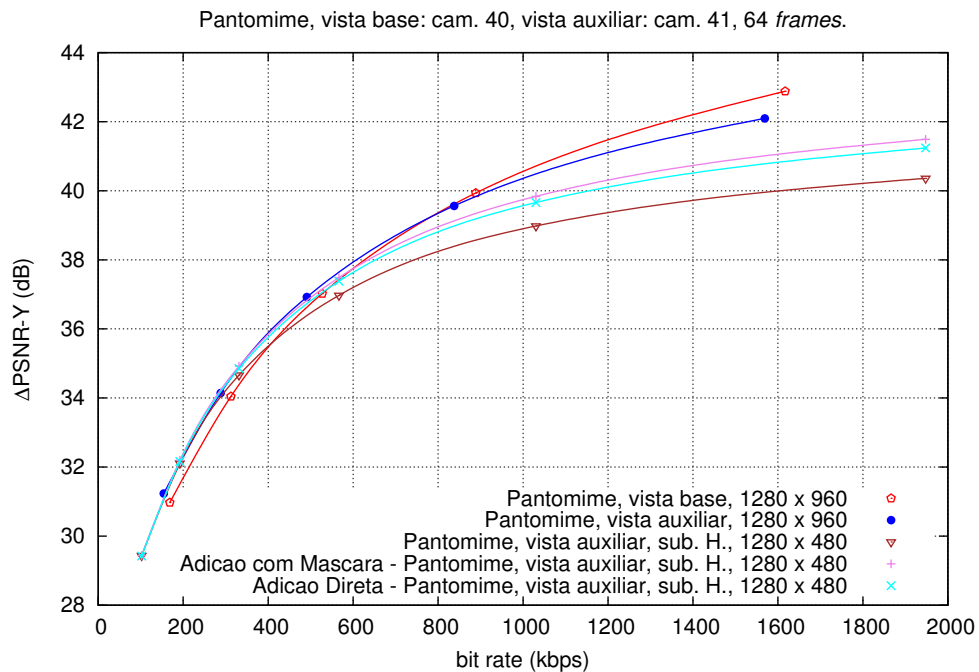


Figura A.85: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Pantomime*.

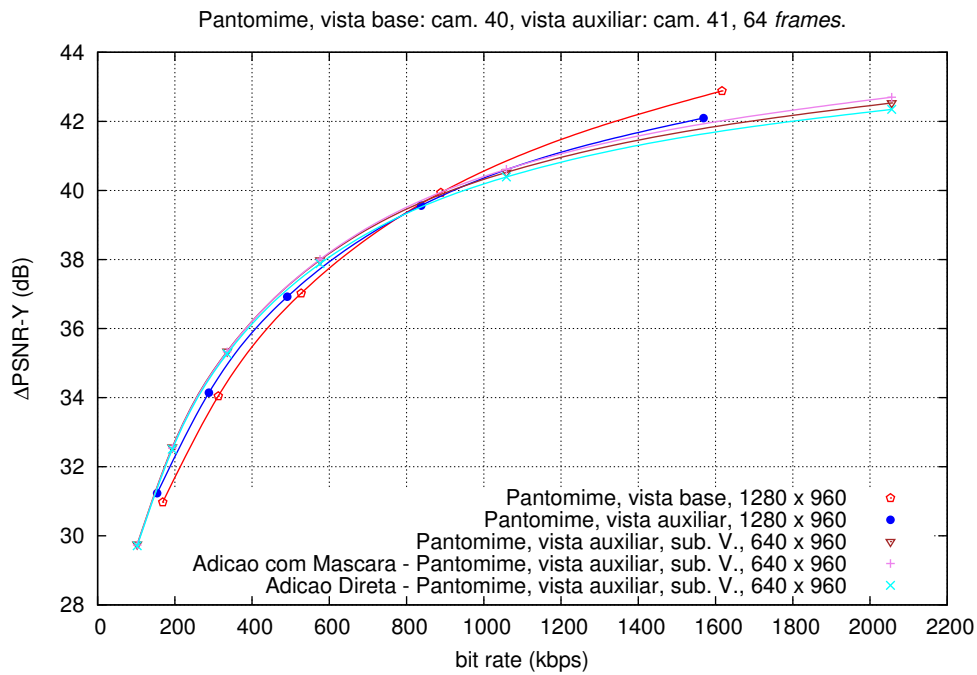


Figura A.86: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência *Pantomime*.

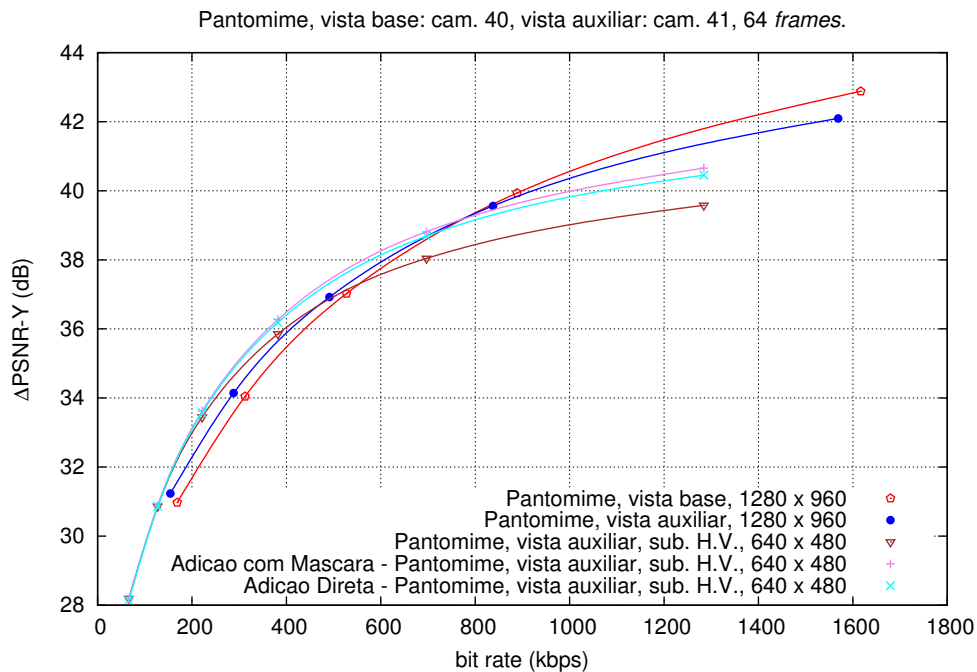


Figura A.87: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência *Pantomime*.

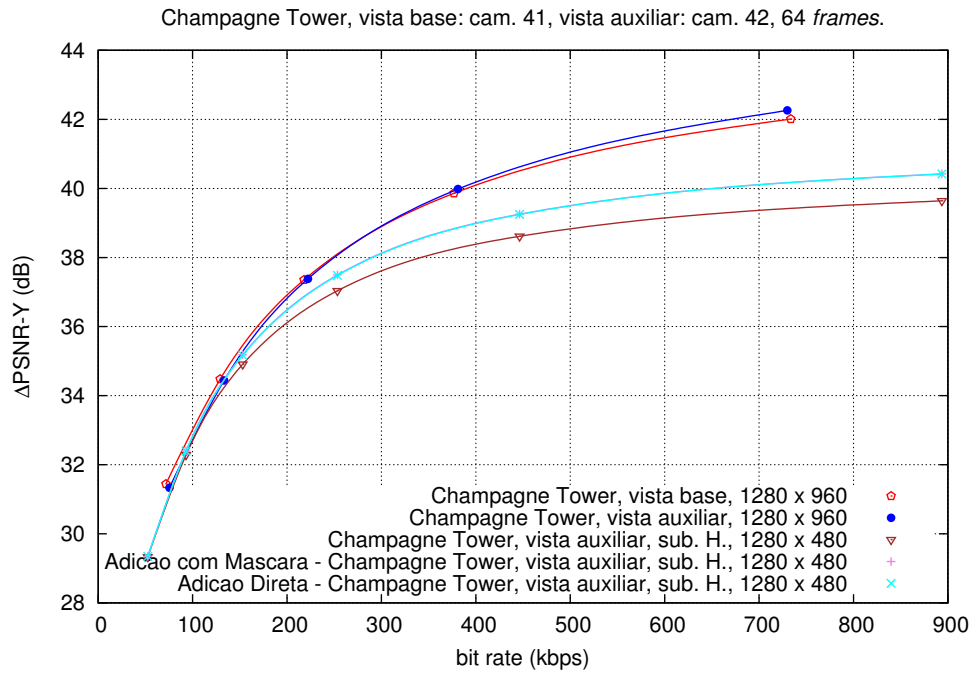


Figura A.88: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Champagne Tower*.

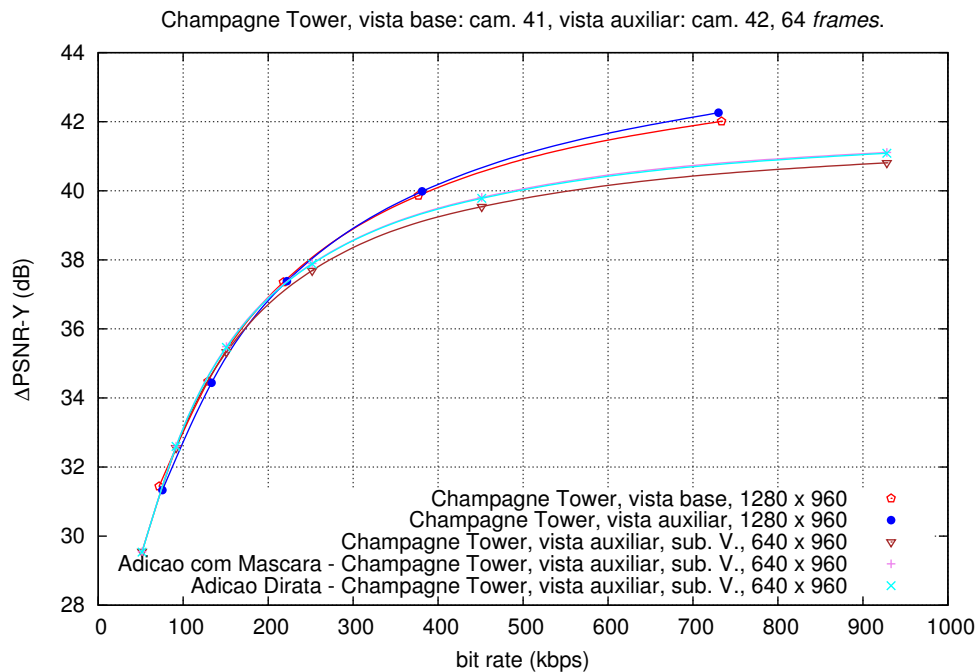


Figura A.89: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência *Champagne Tower*.

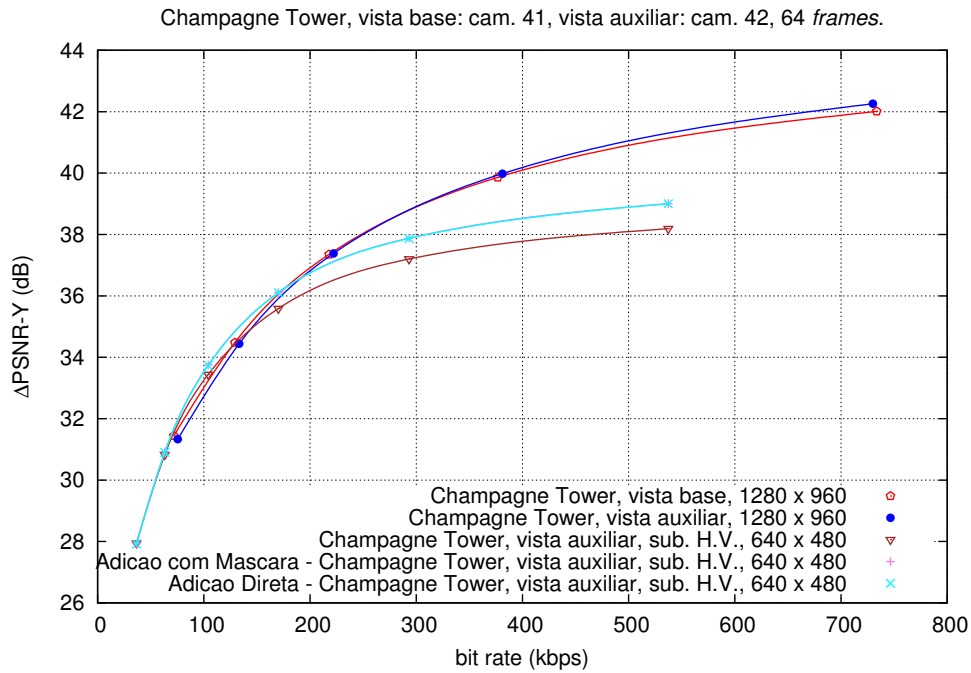


Figura A.90: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência *Champagne Tower*.

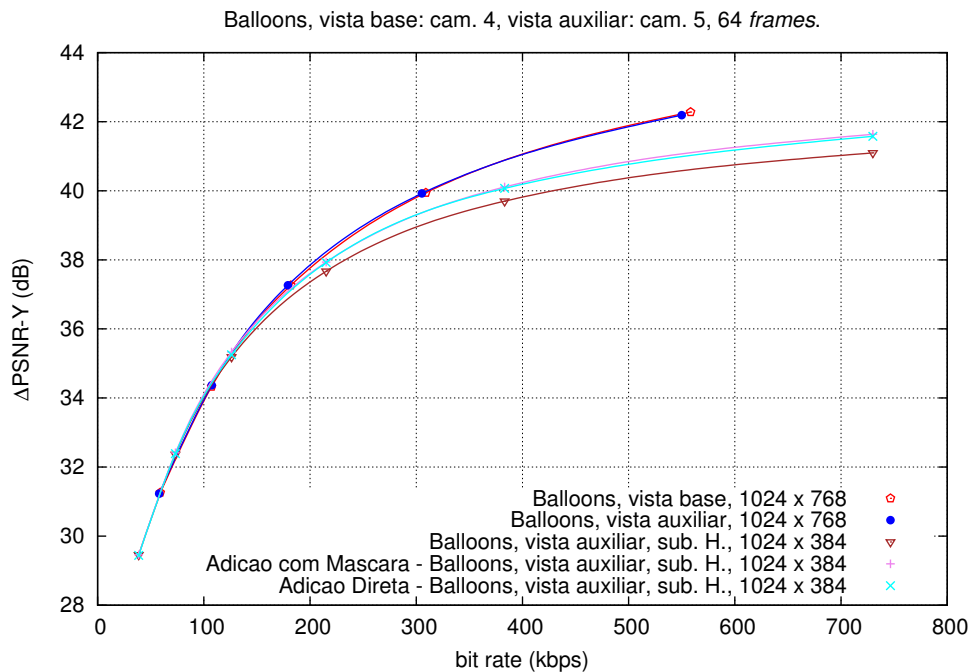


Figura A.91: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Balloons*.

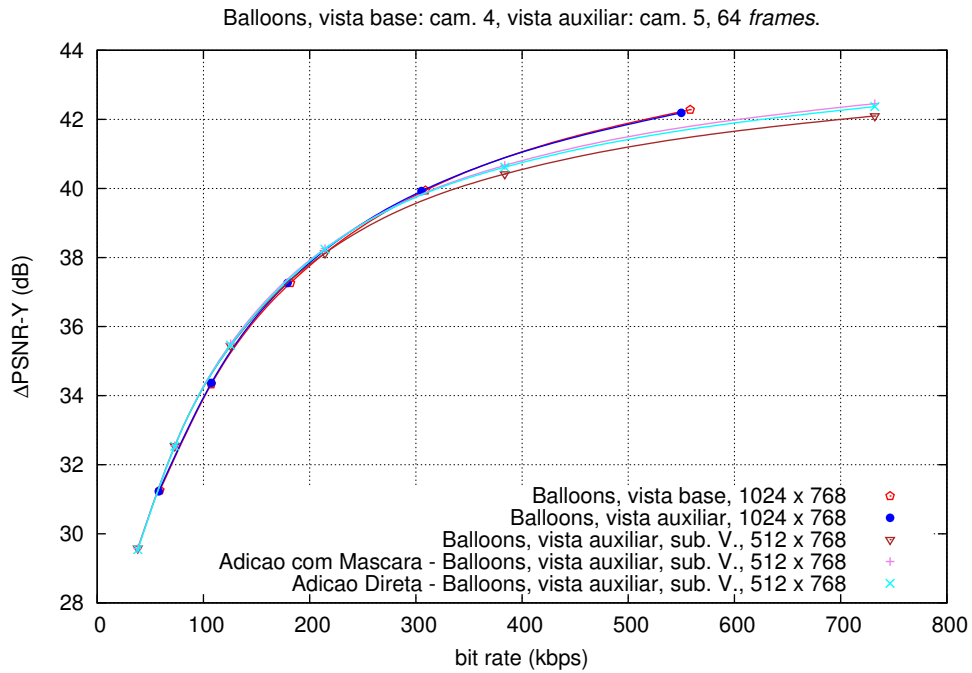


Figura A.92: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V , para a sequência *Balloons*.

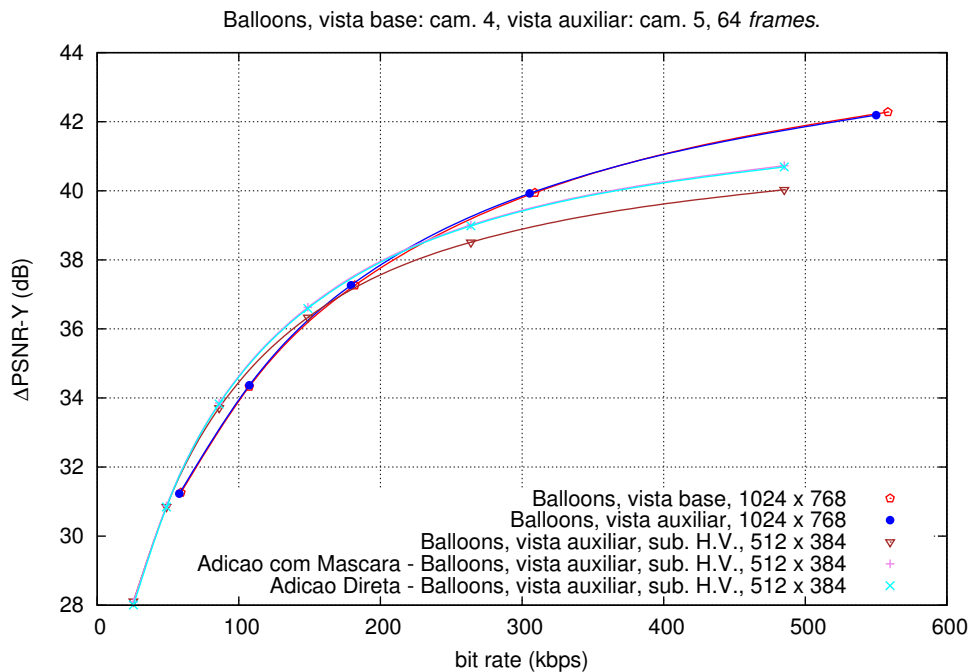


Figura A.93: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V , para a sequência *Balloons*.

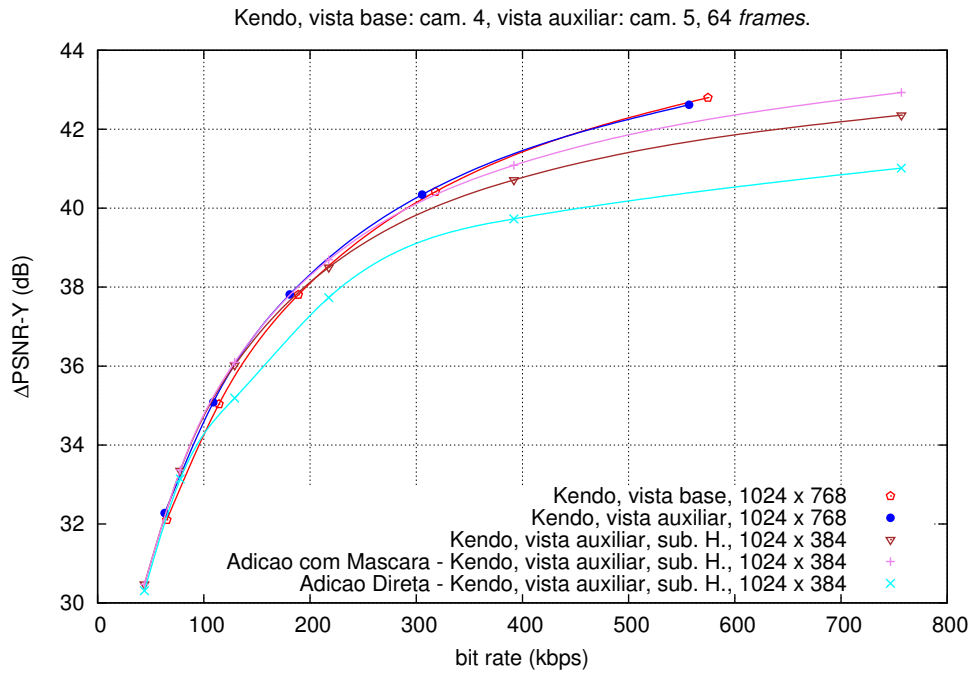


Figura A.94: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Kendo*.

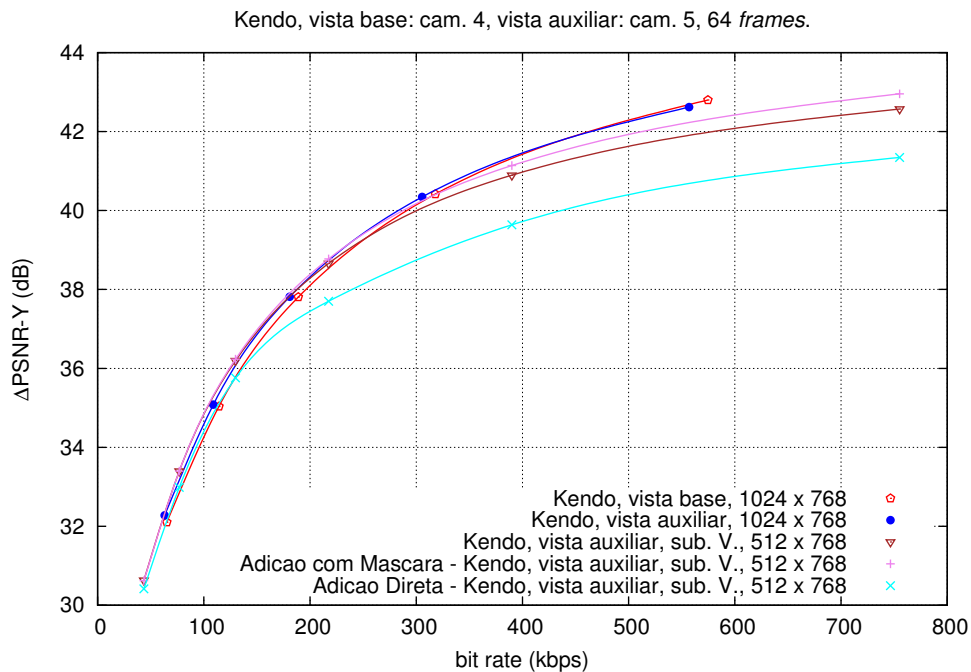


Figura A.95: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência *Kendo*.

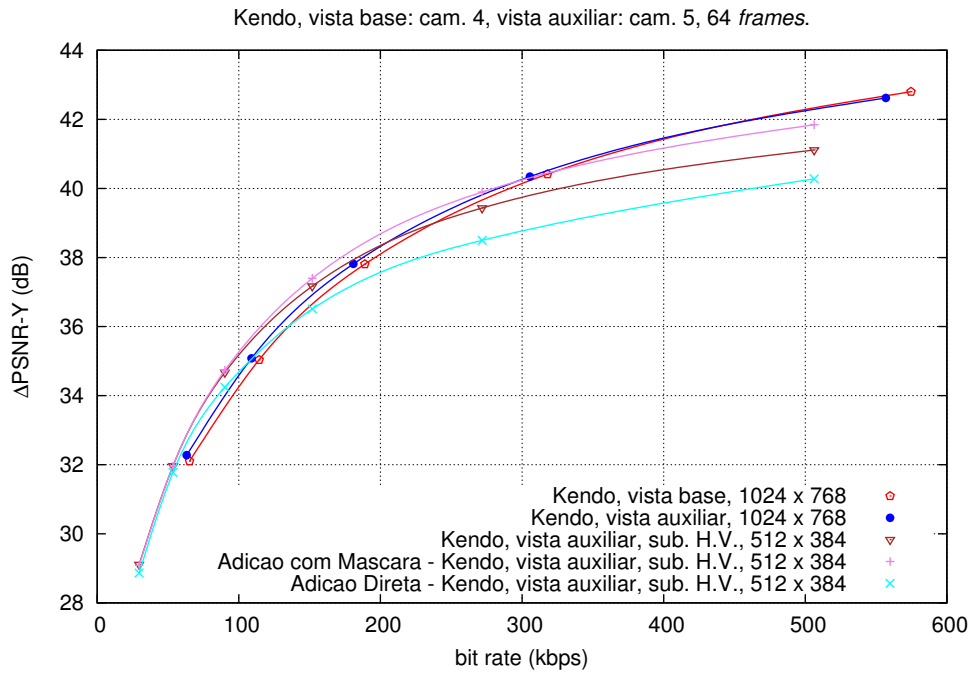


Figura A.96: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência *Kendo*.

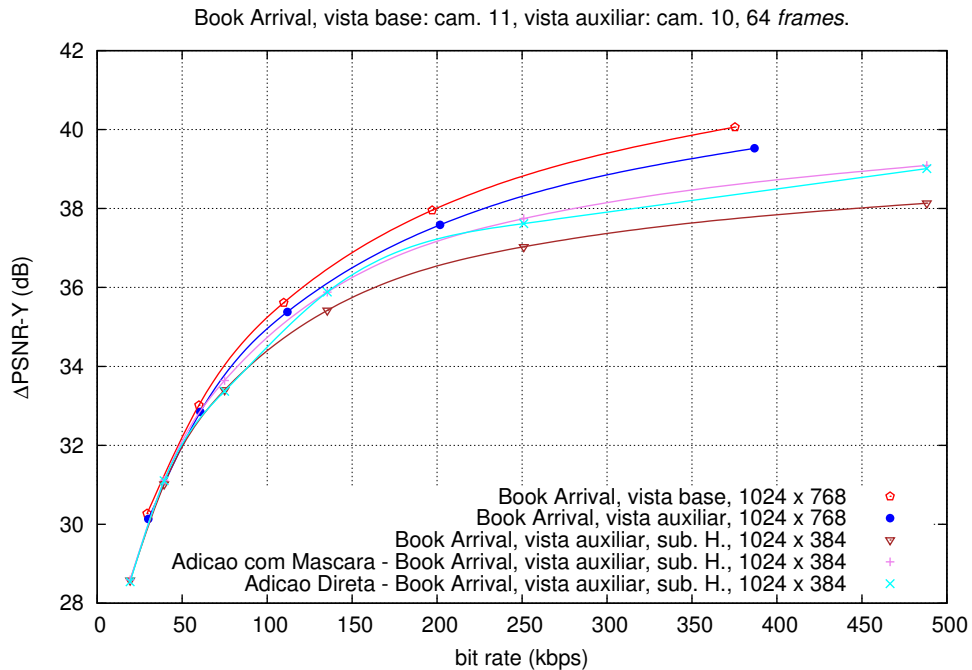


Figura A.97: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Book Arrival*.

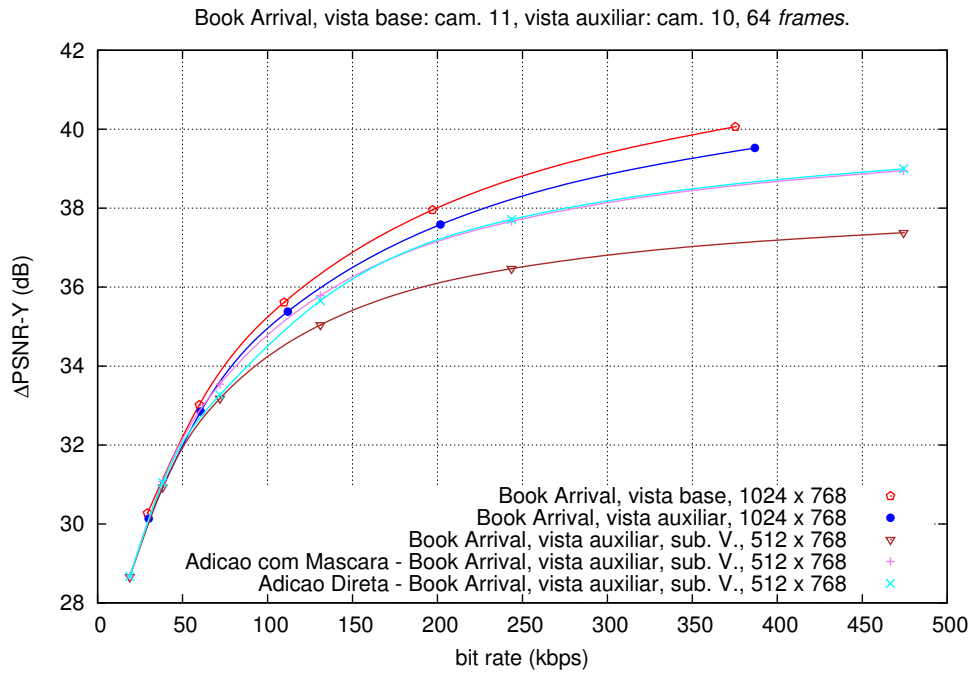


Figura A.98: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência *Book Arrival*.

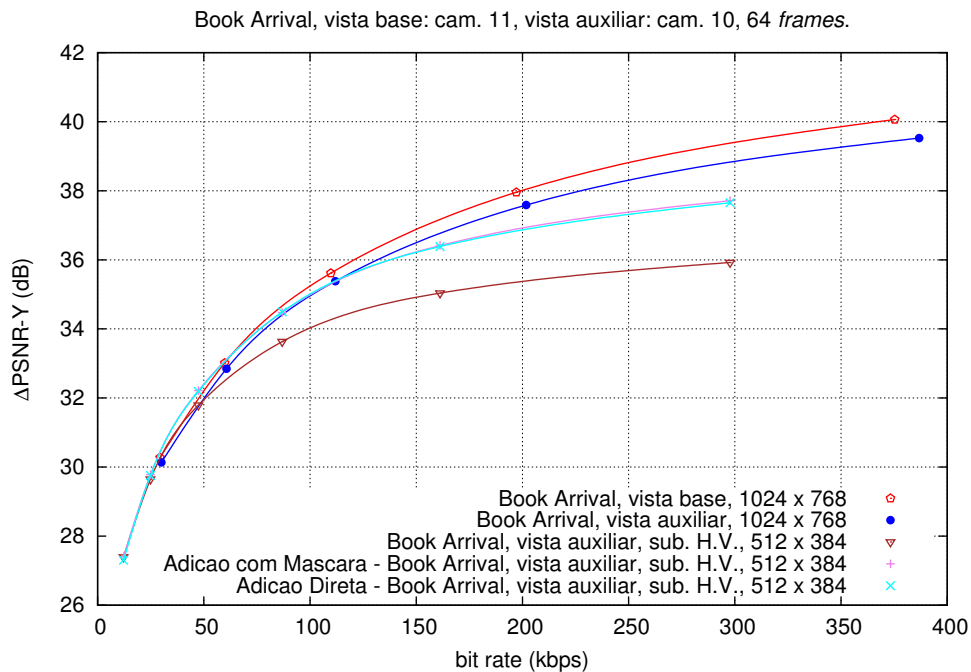


Figura A.99: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência *Book Arrival*.

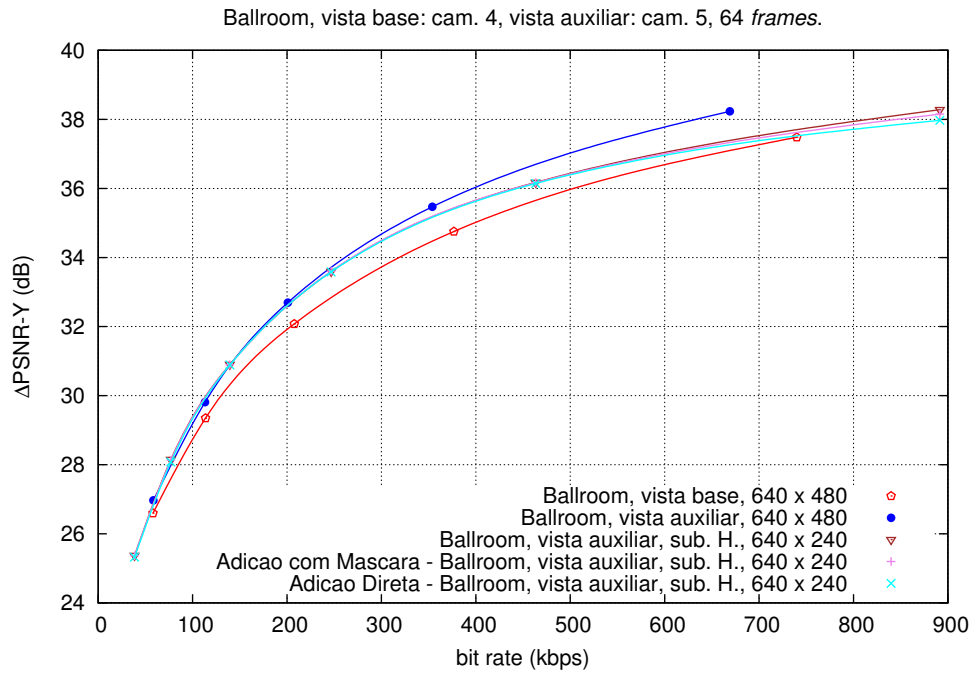


Figura A.100: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Ballroom*.

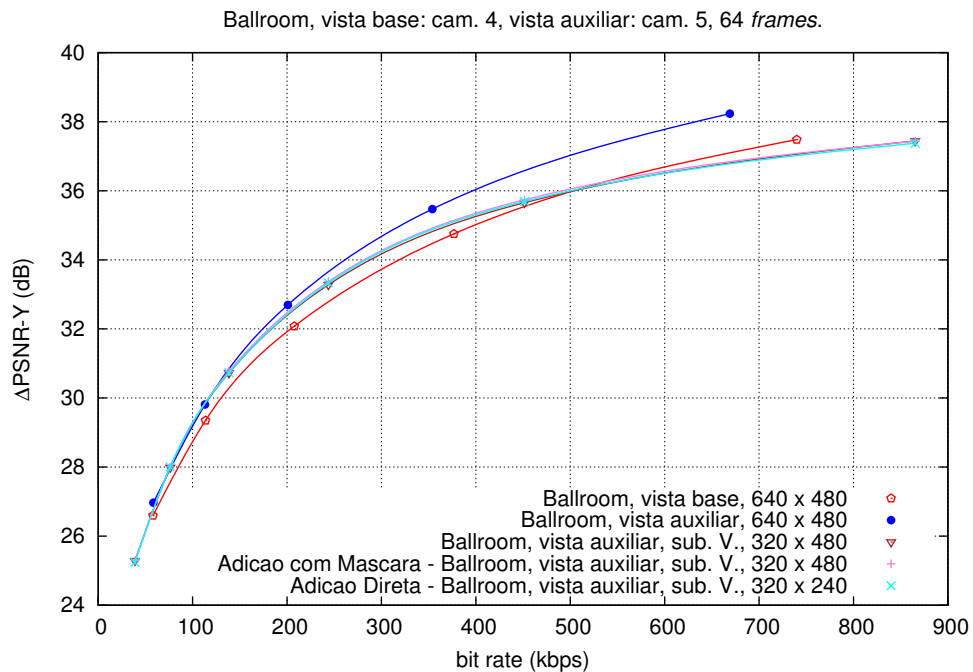


Figura A.101: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V, para a sequência *Ballroom*.

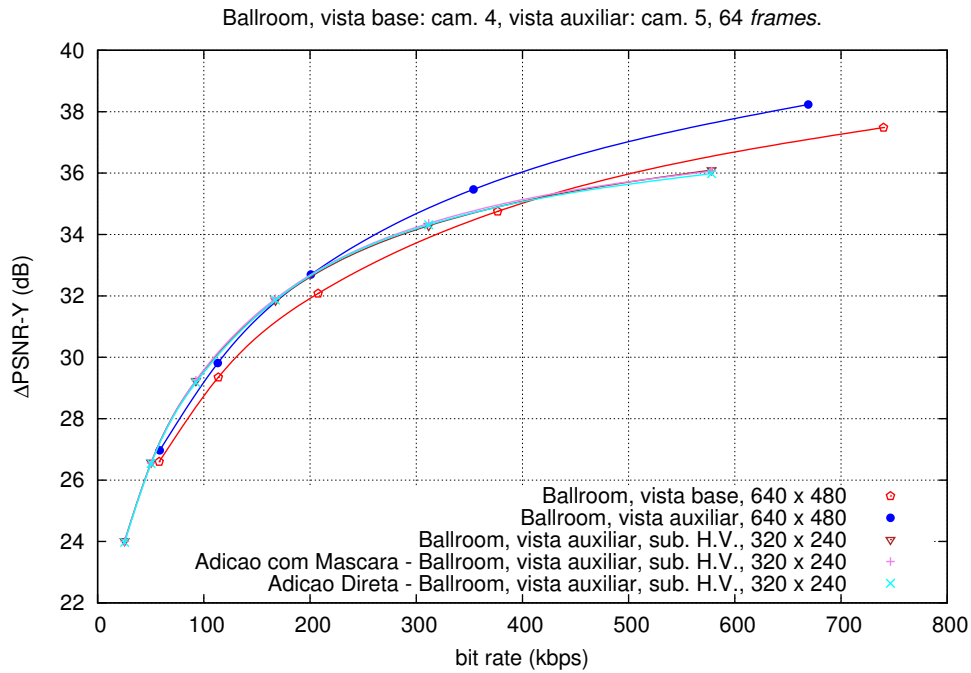


Figura A.102: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V, para a sequência *Ballroom*.

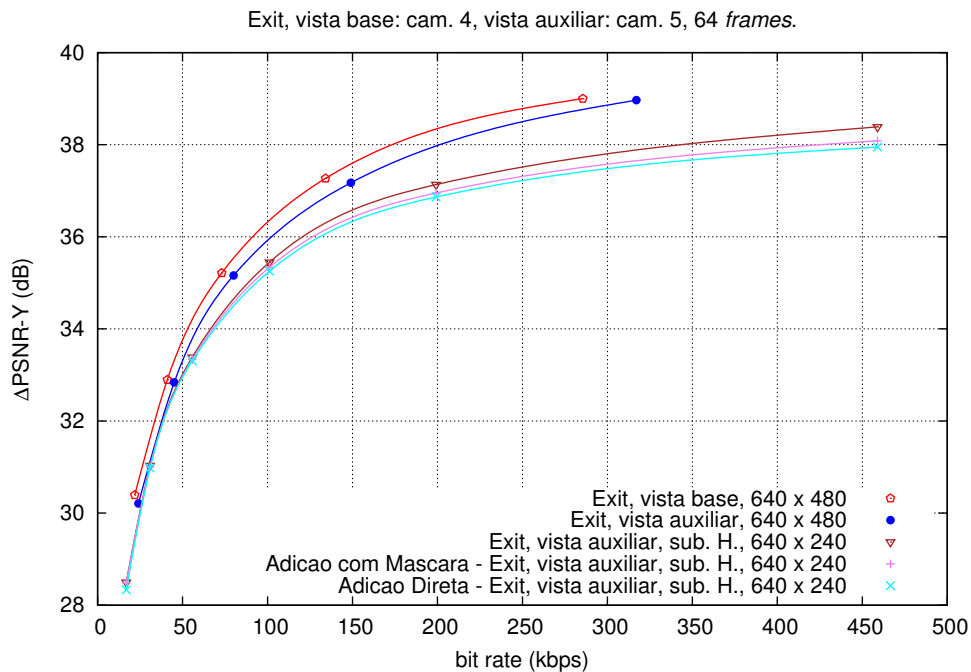


Figura A.103: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H, para a sequência *Exit*.

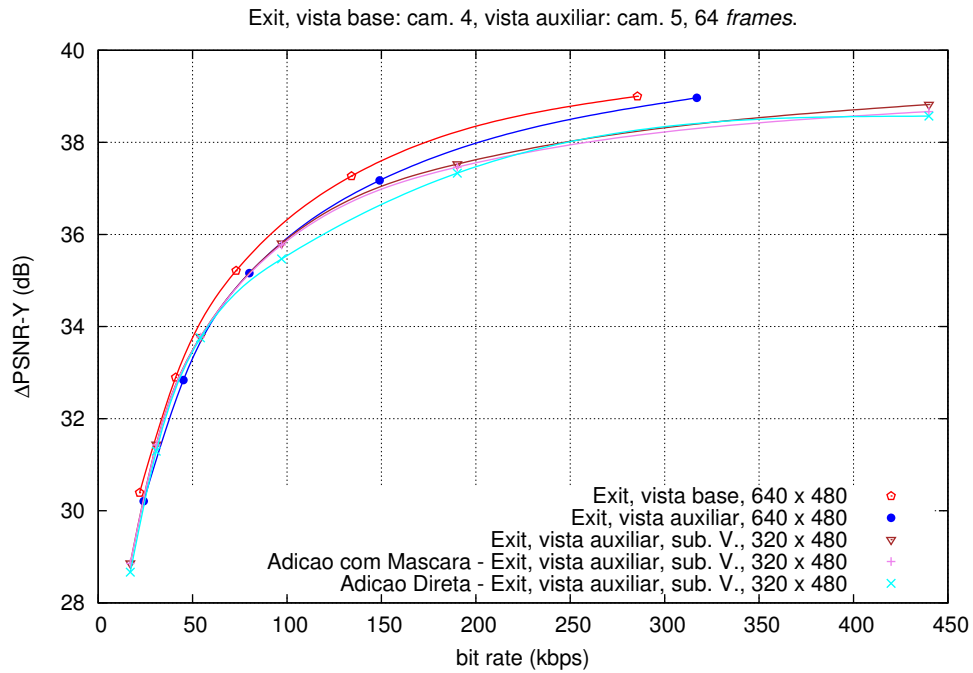


Figura A.104: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar V , para a sequência *Exit*.

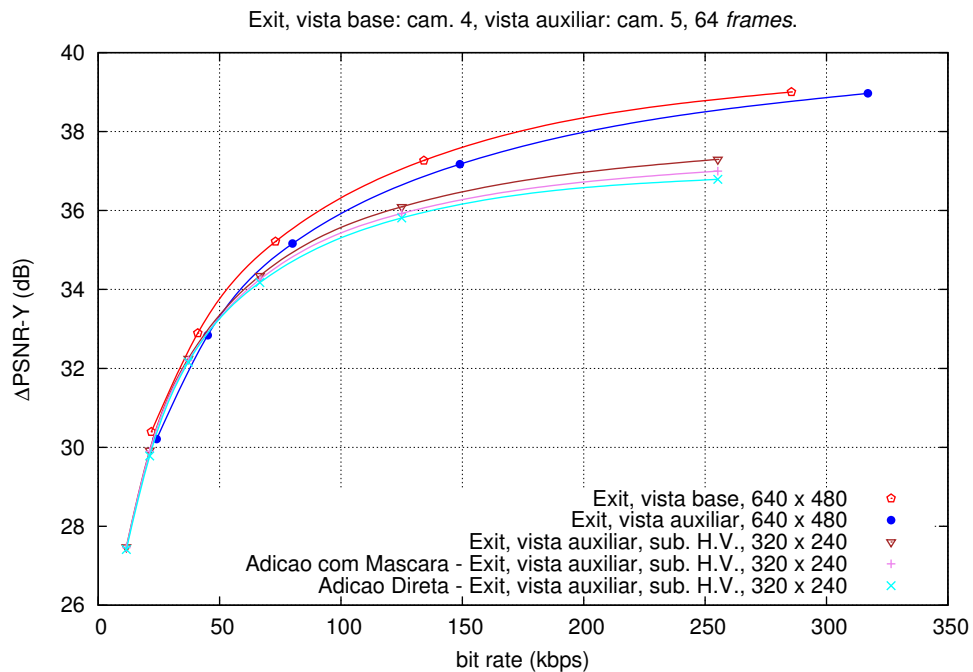


Figura A.105: Comparação entre adição direta e com base na máscara, no método da malha com a vista auxiliar H/V , para a sequência *Exit*.

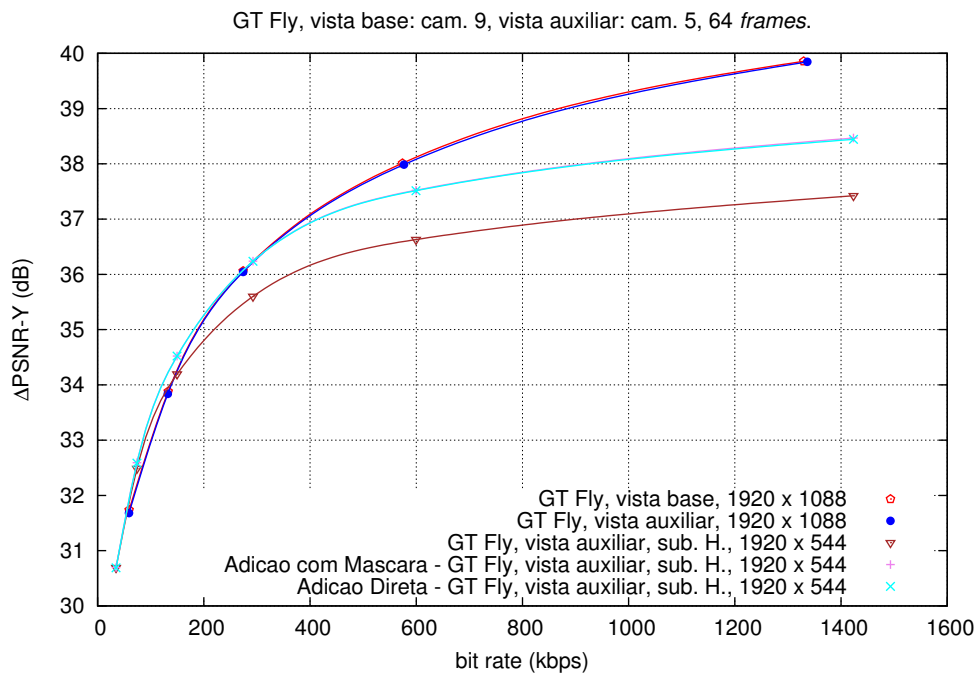


Figura A.106: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *GT Fly*.

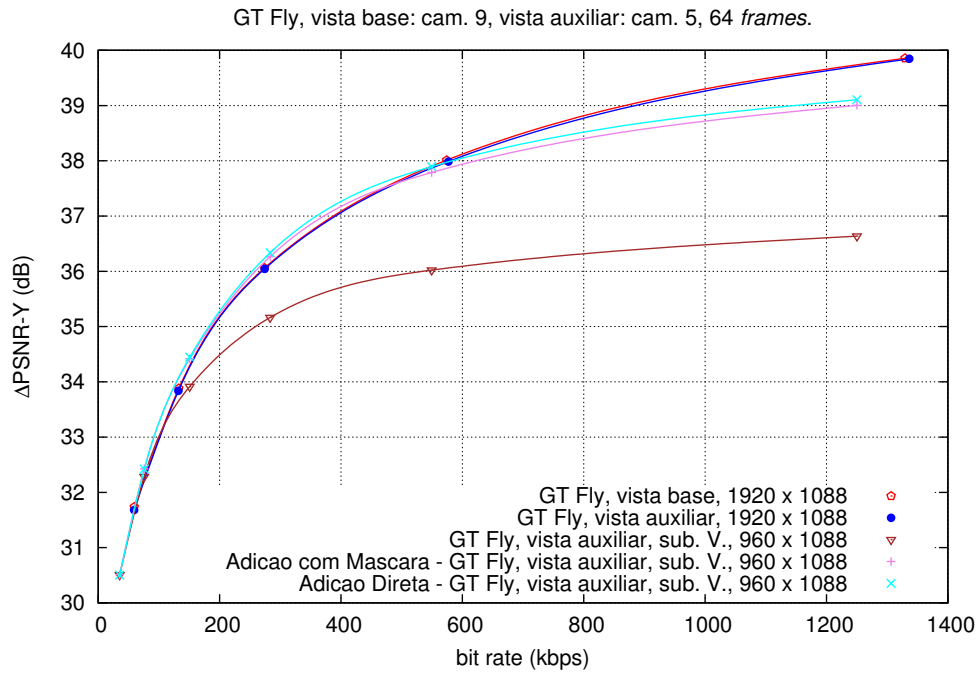


Figura A.107: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *GT Fly*.

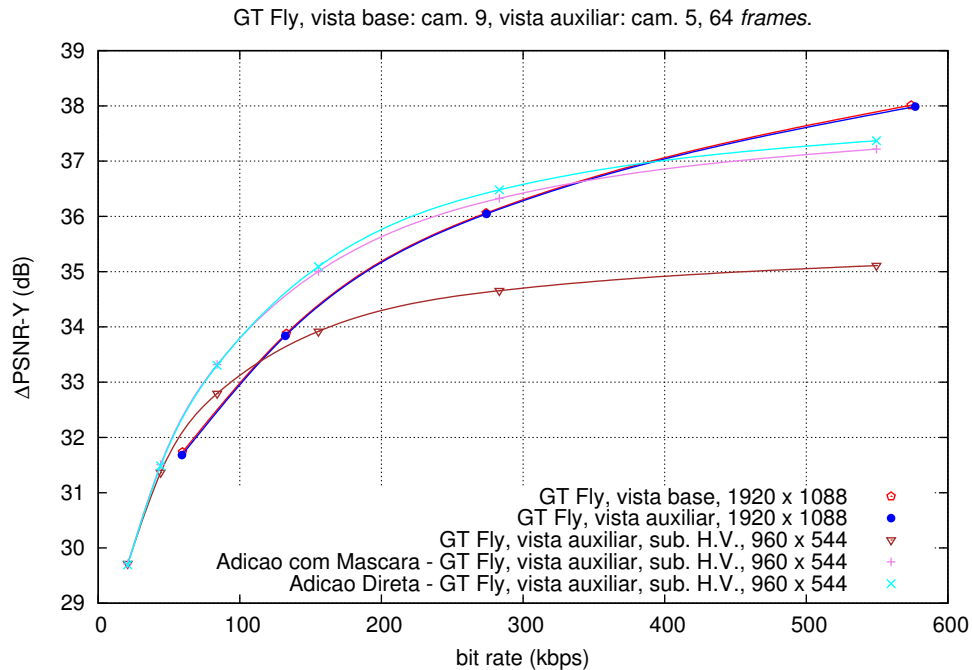


Figura A.108: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *GT Fly*.

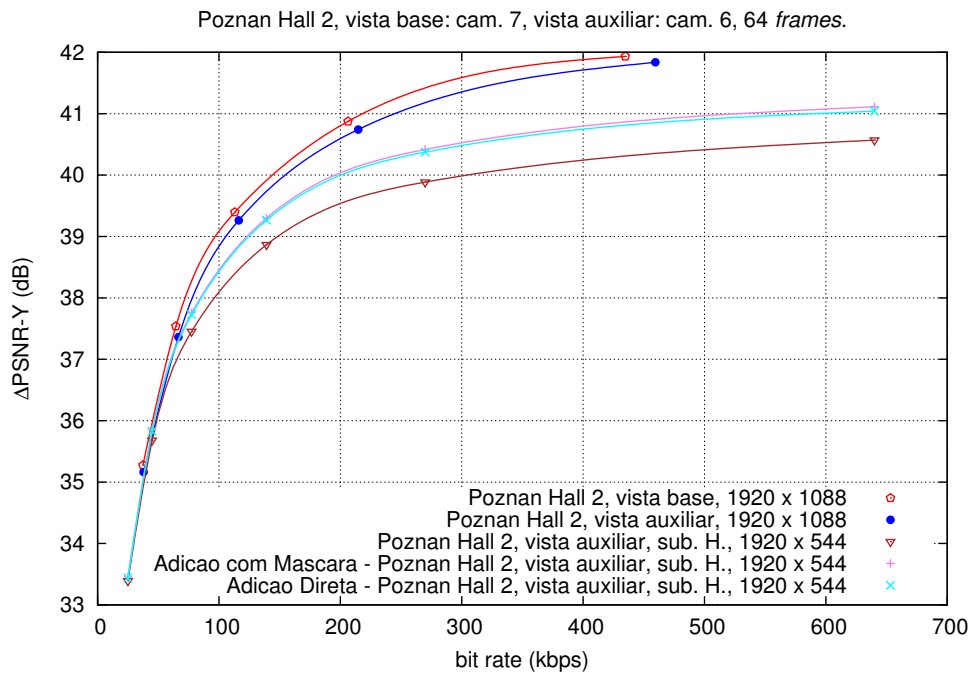


Figura A.109: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Poznan Hall 2*.

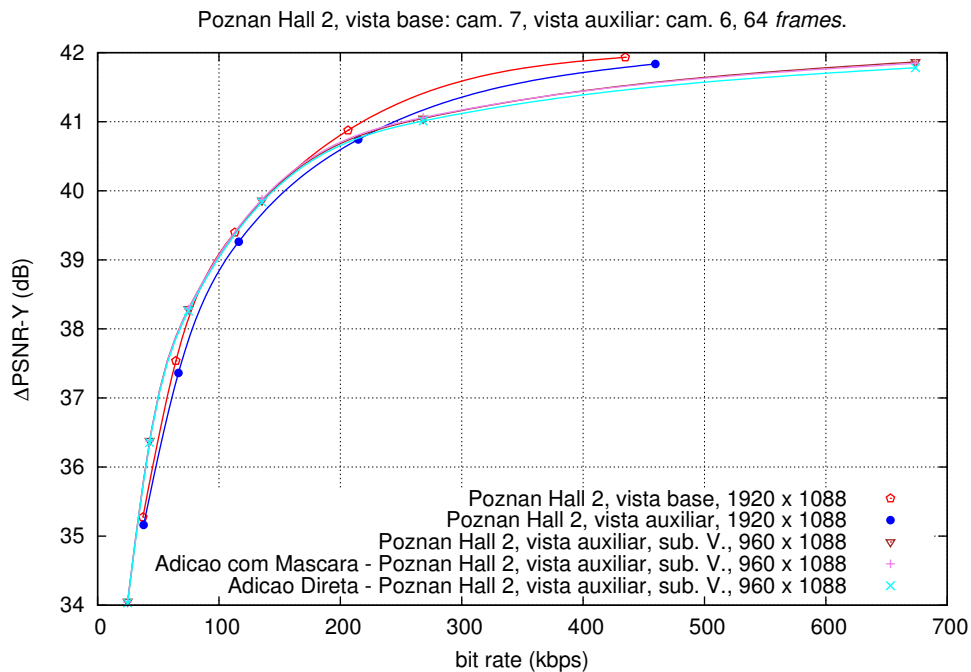


Figura A.110: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *Poznan Hall 2*.

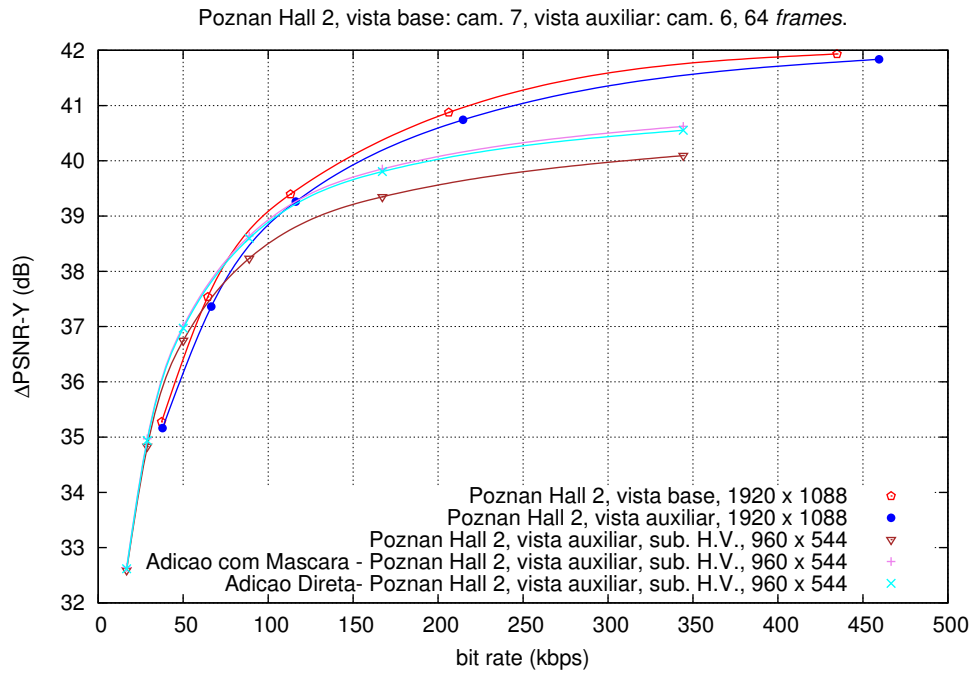


Figura A.111: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *Poznan Hall 2*.

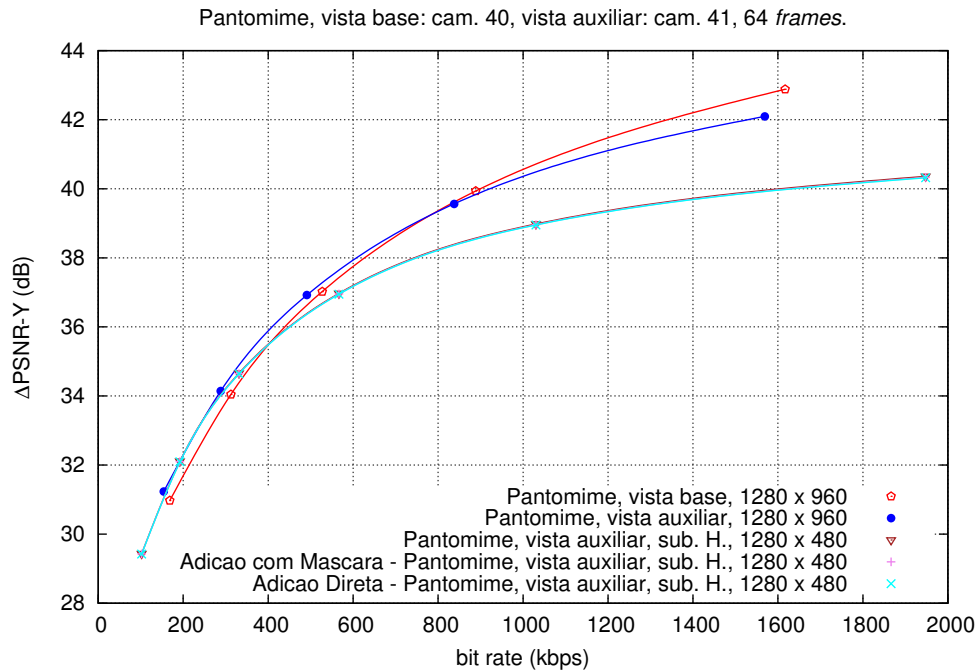


Figura A.112: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Pantomime*.

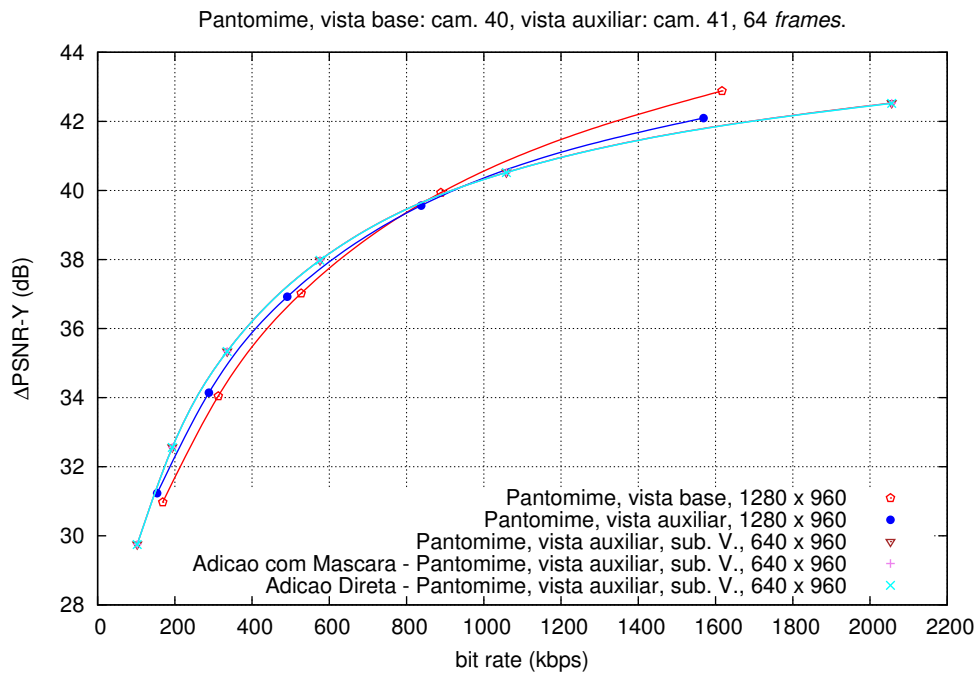


Figura A.113: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *Pantomime*.

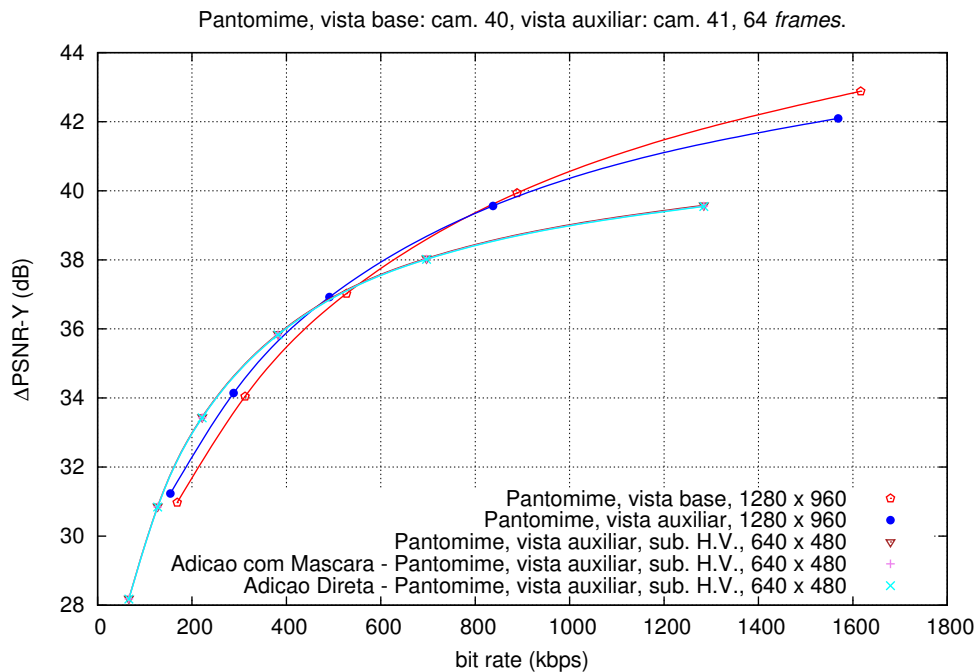


Figura A.114: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *Pantomime*.

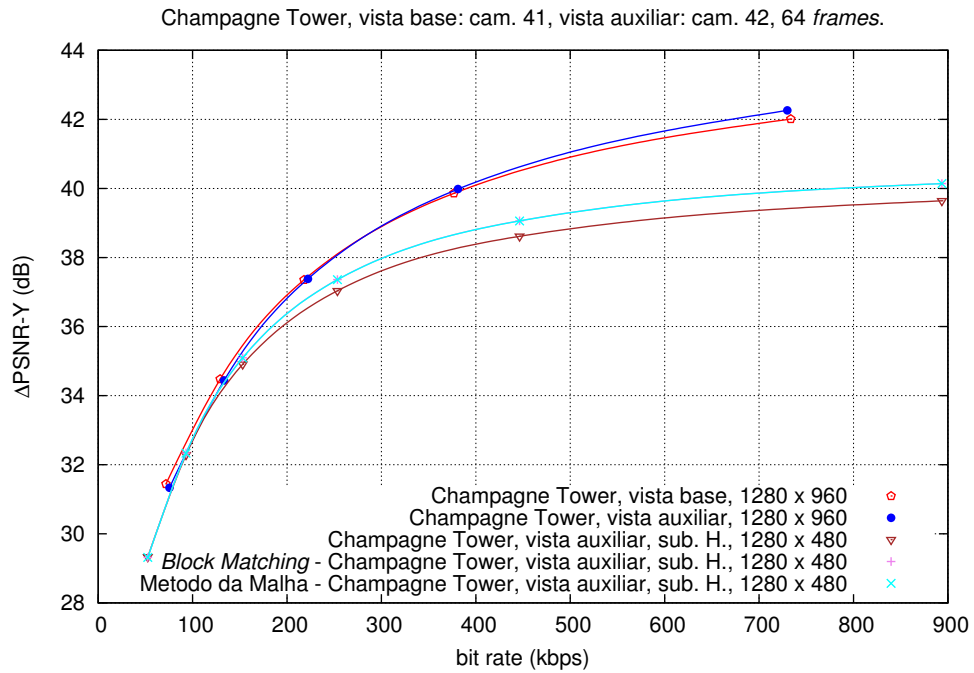


Figura A.115: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Champagne Tower*.

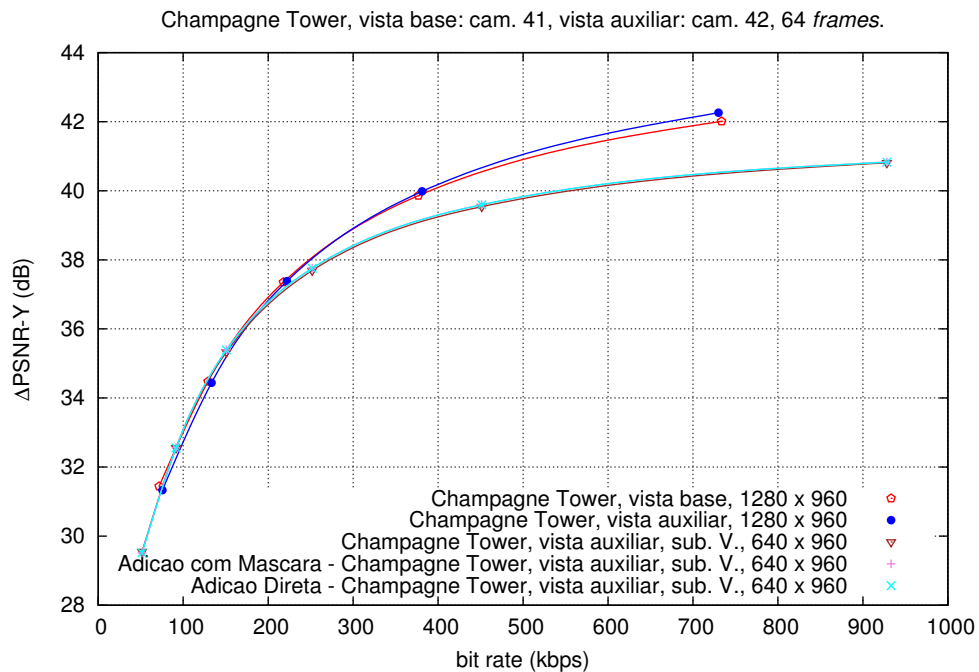


Figura A.116: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *Champagne Tower*.

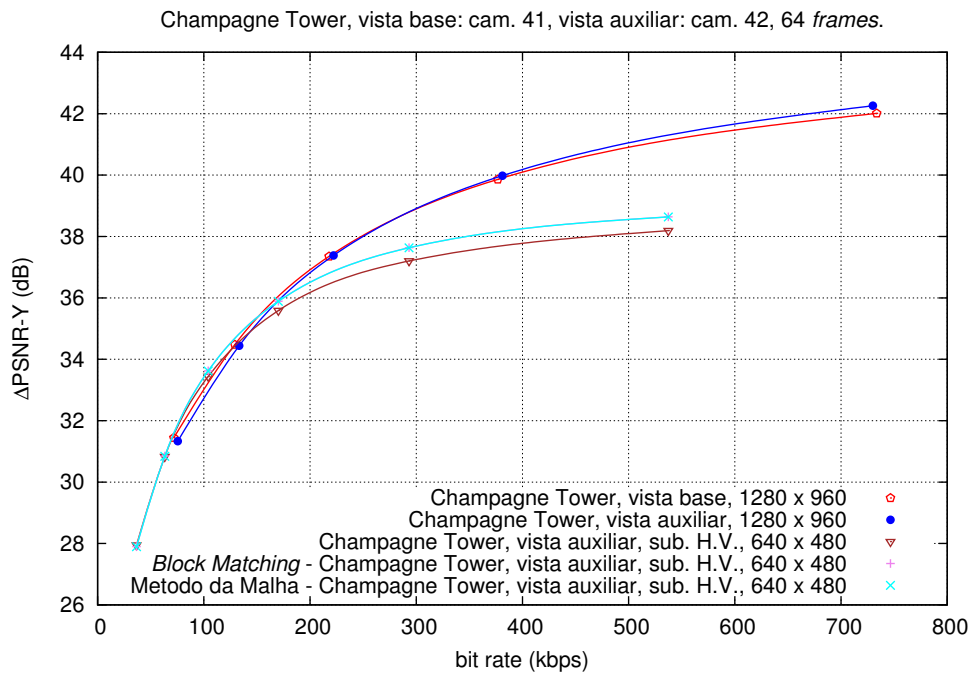


Figura A.117: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *Champagne Tower*.

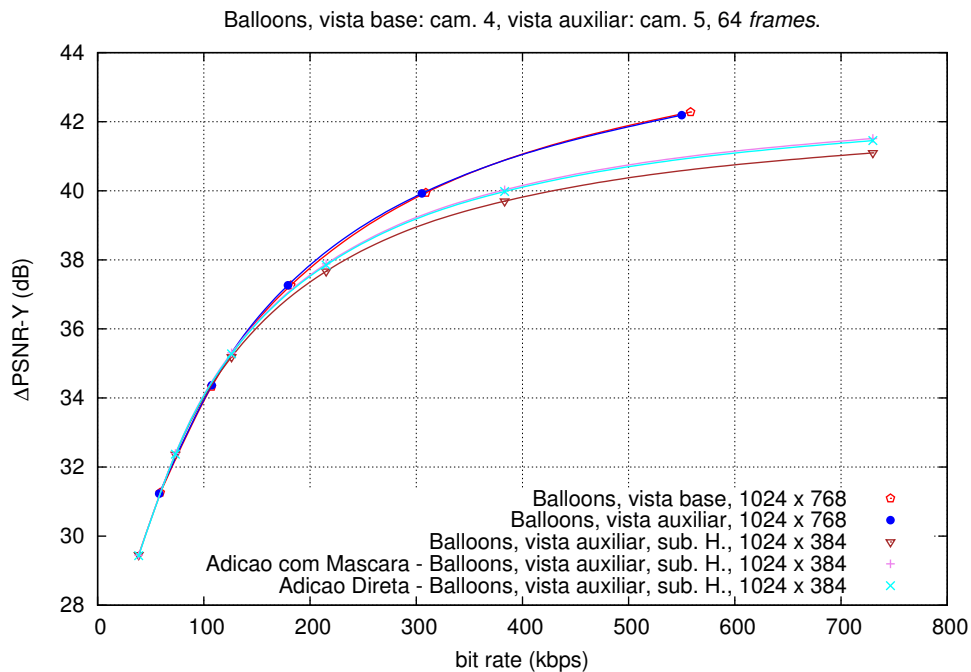


Figura A.118: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Balloons*.

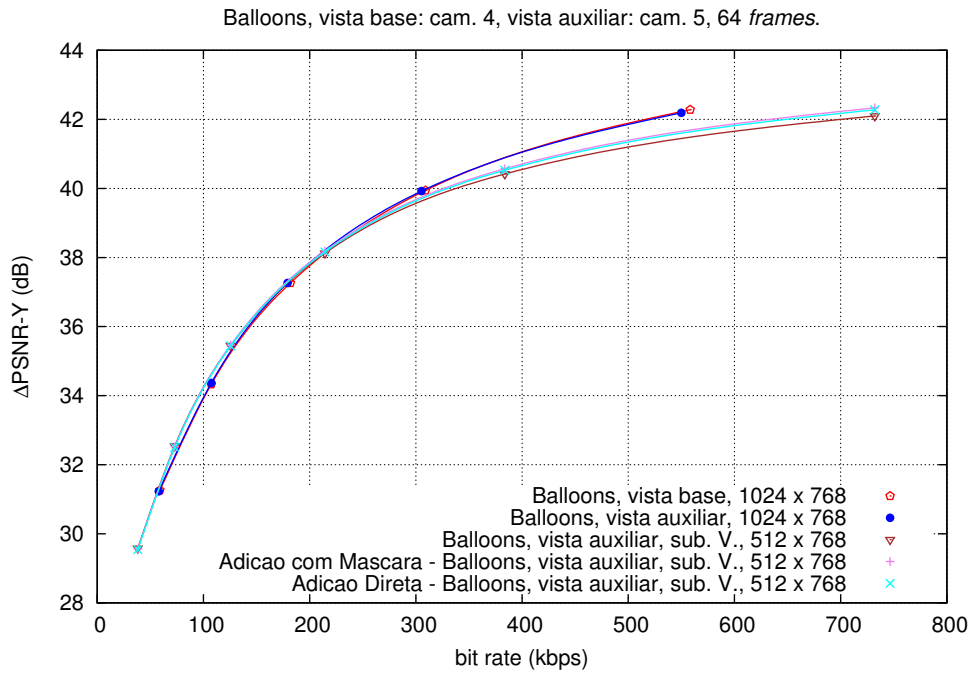


Figura A.119: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V , para a sequência *Balloons*.

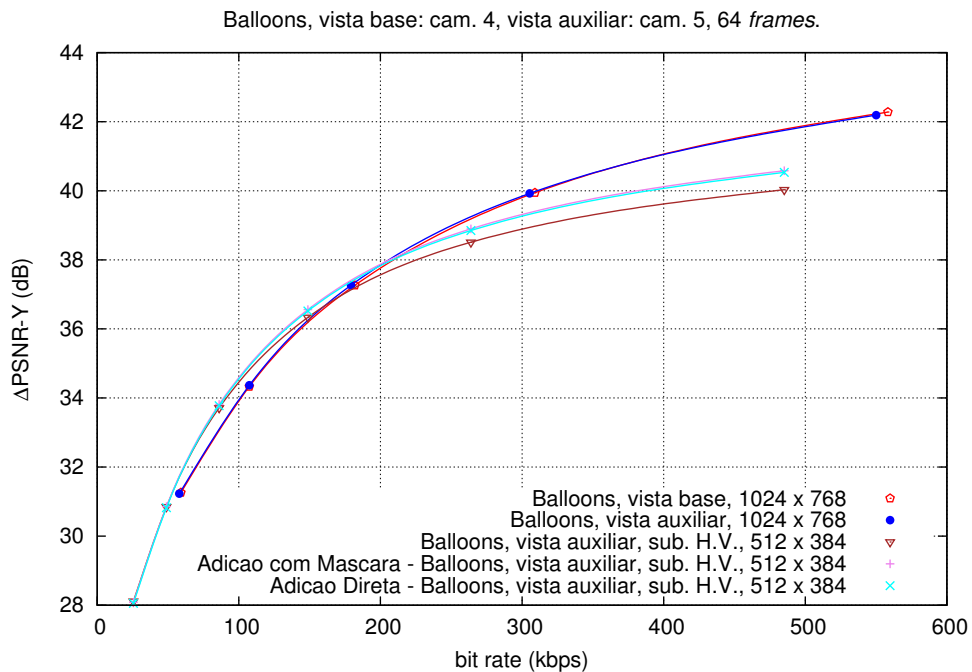


Figura A.120: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V , para a sequência *Balloons*.

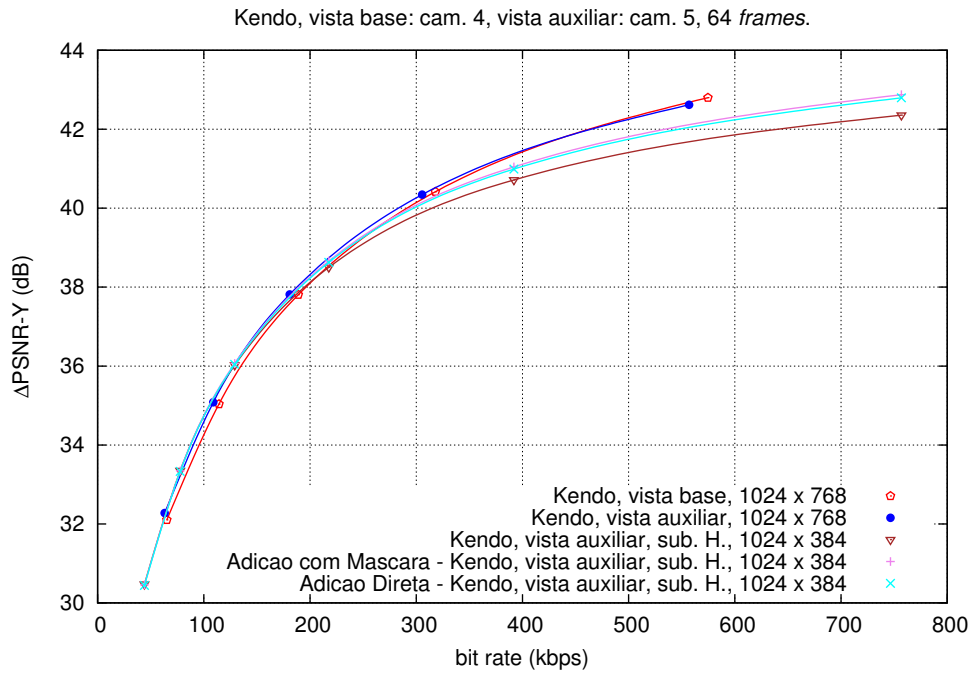


Figura A.121: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Kendo*.

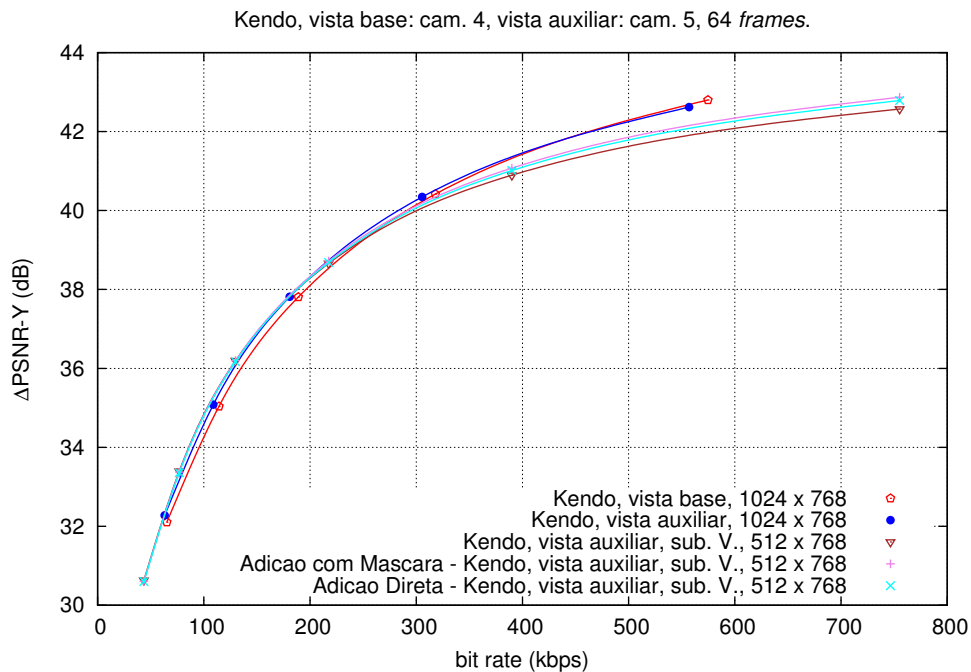


Figura A.122: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *Kendo*.

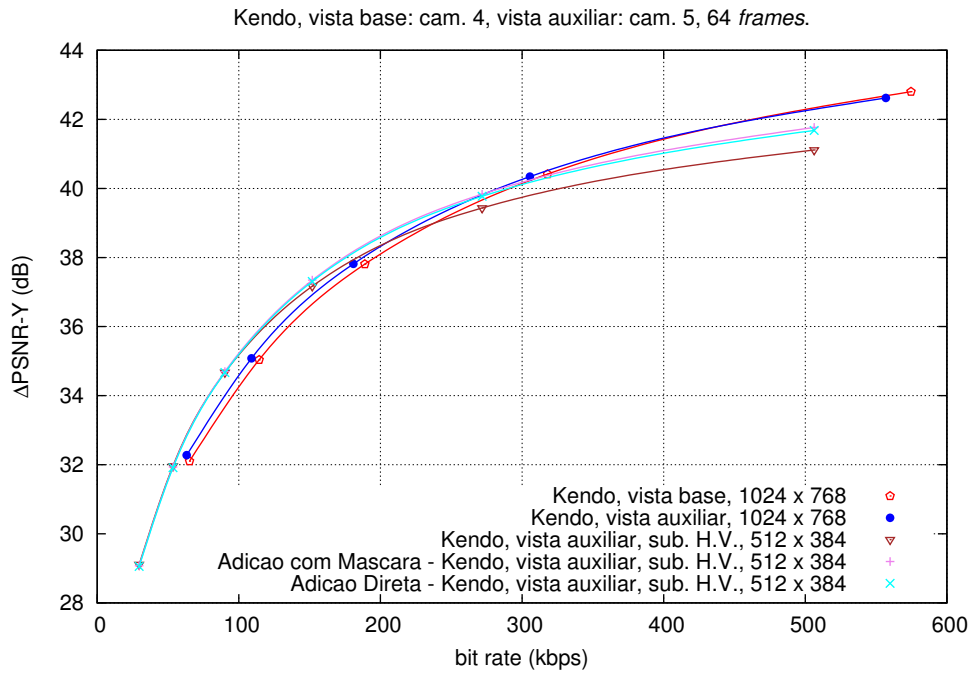


Figura A.123: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *Kendo*.

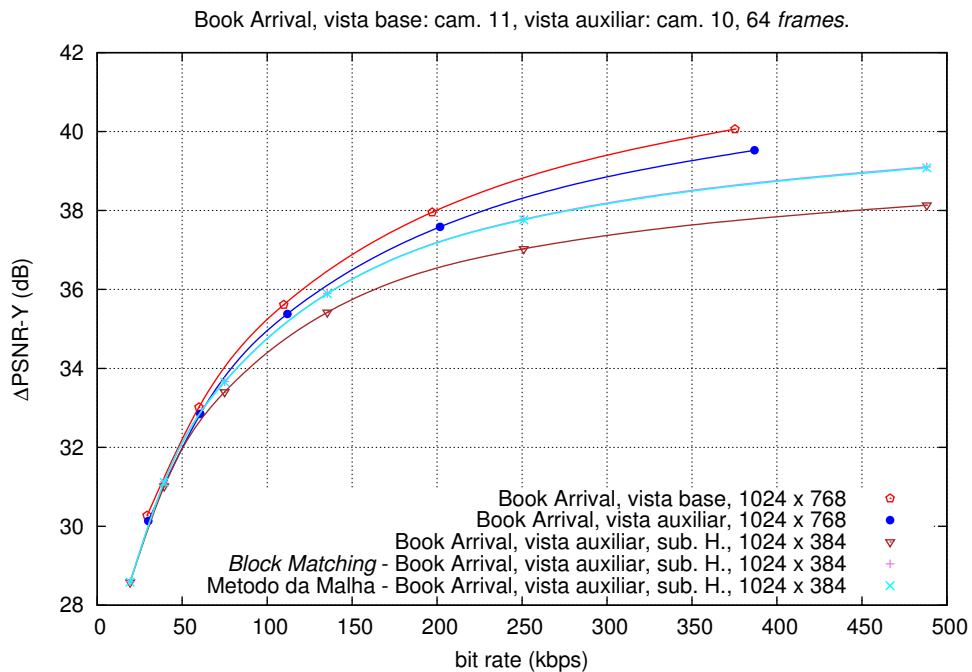


Figura A.124: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Book Arrival*.

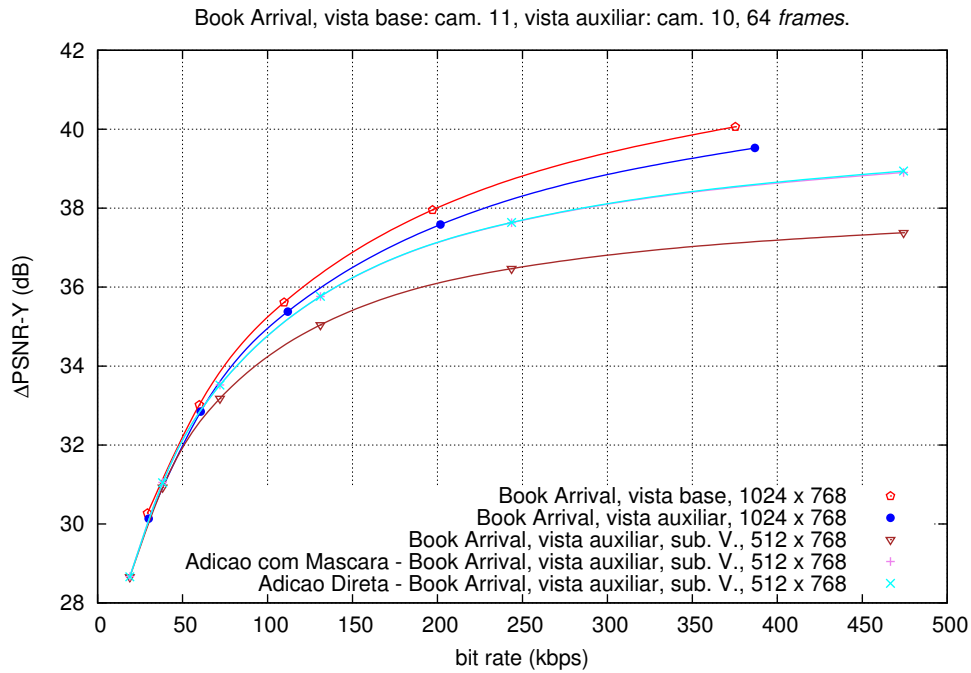


Figura A.125: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *Book Arrival*.

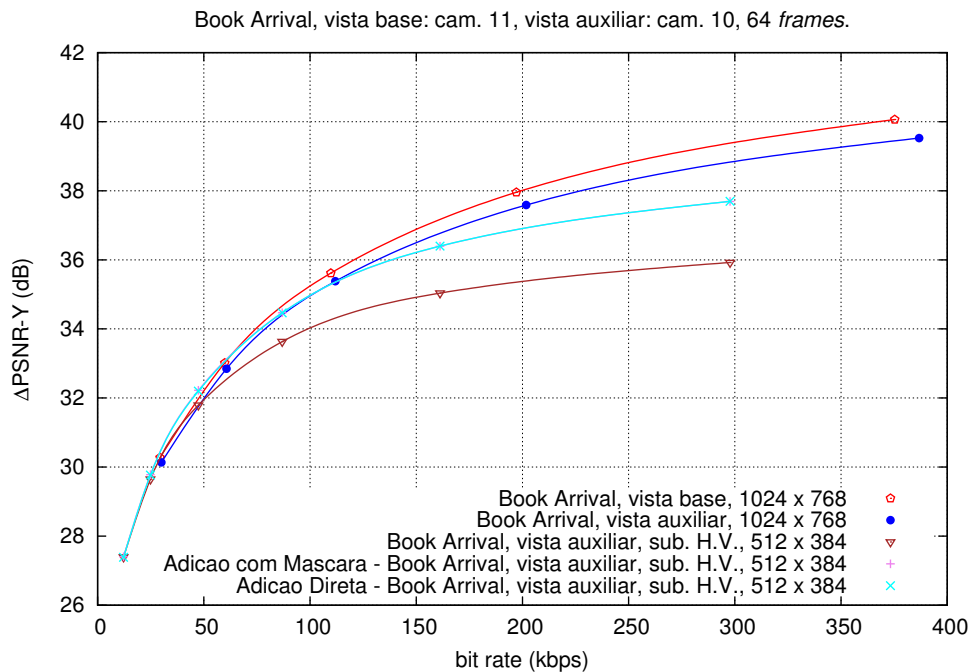


Figura A.126: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *Book Arrival*.

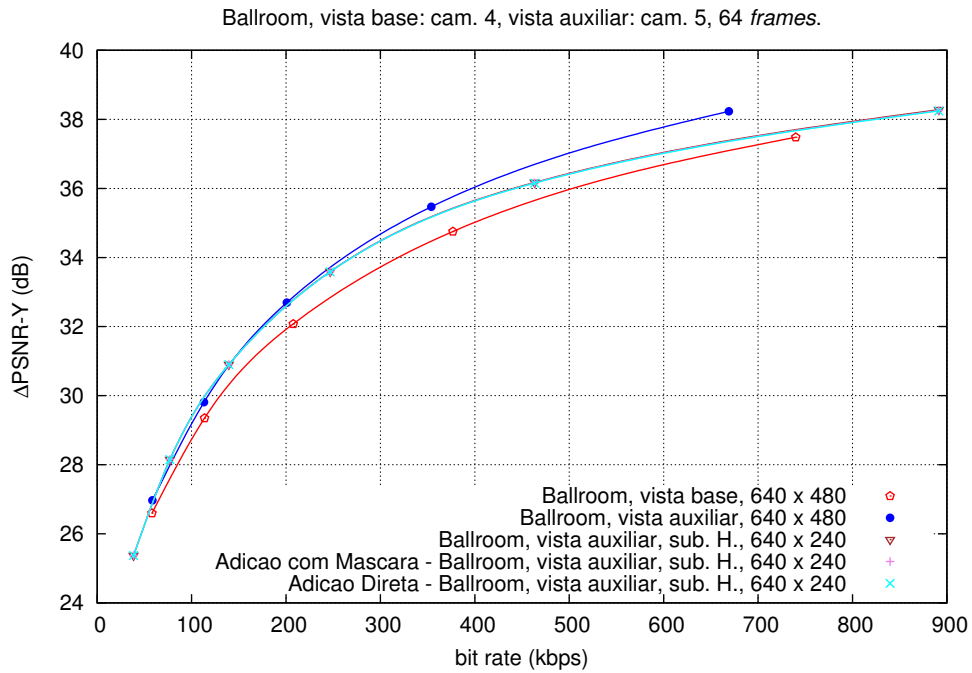


Figura A.127: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Ballroom*.

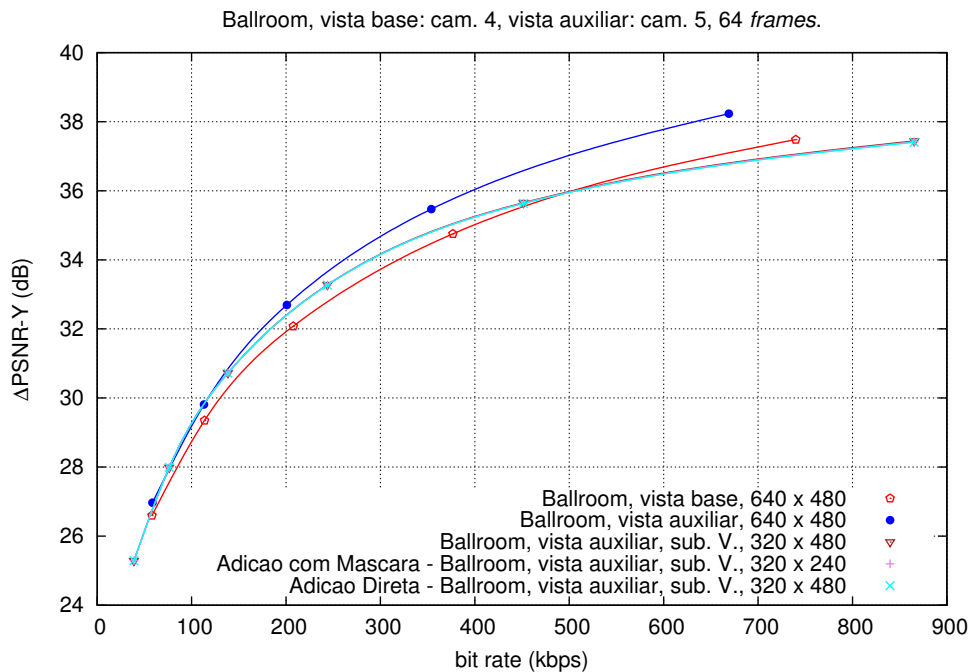


Figura A.128: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V, para a sequência *Ballroom*.

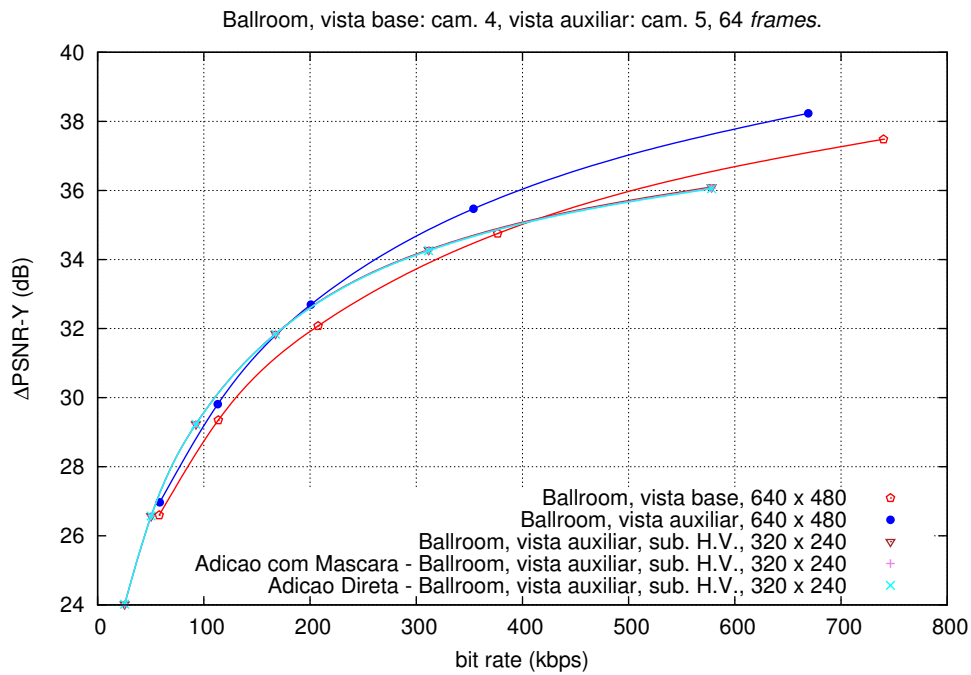


Figura A.129: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V, para a sequência *Ballroom*.

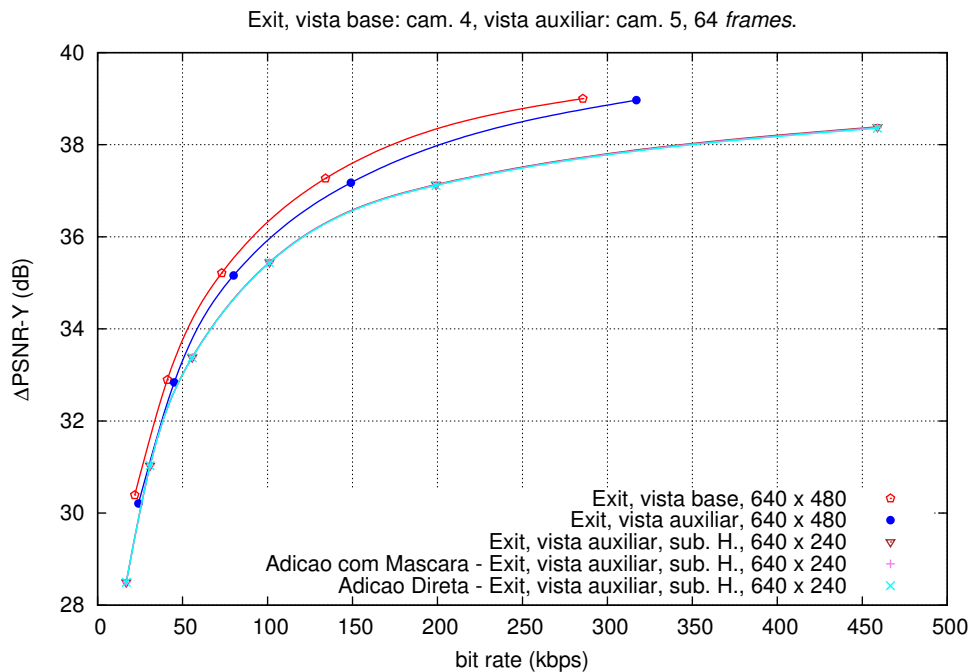


Figura A.130: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H, para a sequência *Exit*.

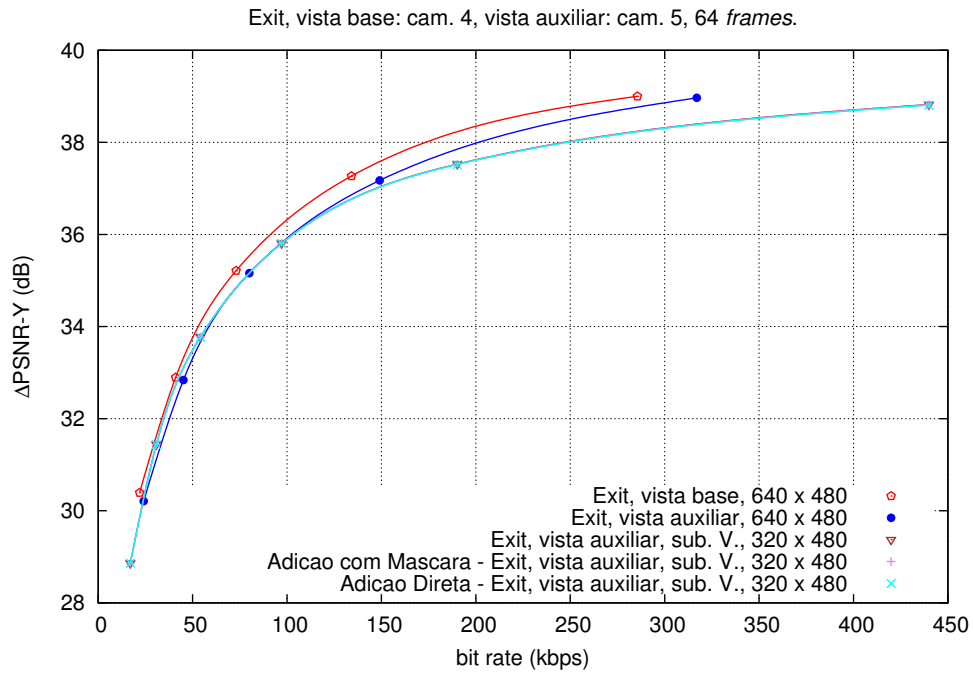


Figura A.131: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar V , para a sequência *Exit*.

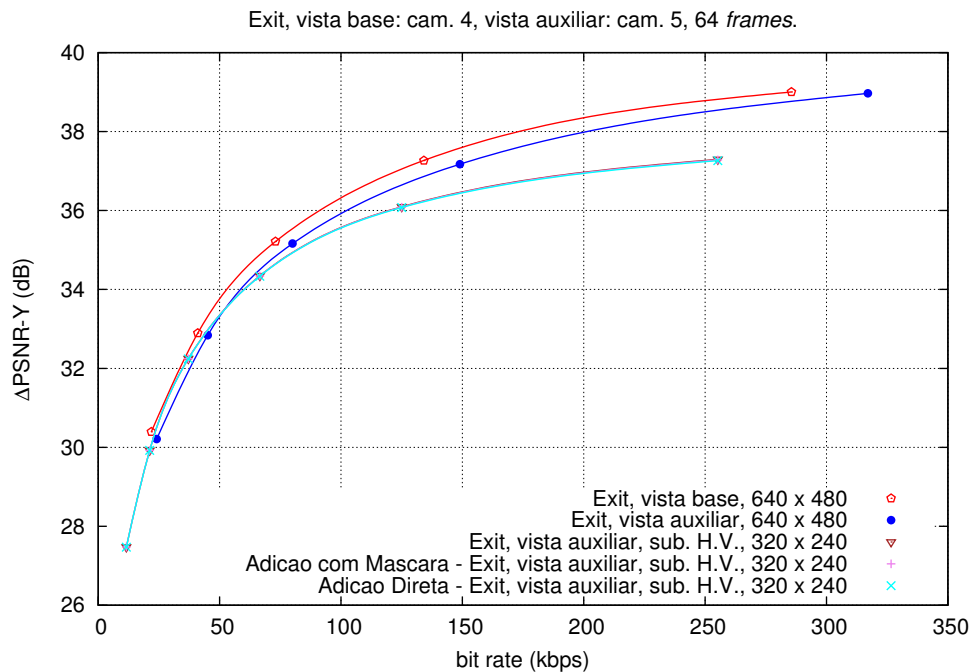


Figura A.132: Comparação entre adição direta e com base na máscara, no método da pesquisa de blocos com a vista auxiliar H/V , para a sequência *Exit*.

A.7 Avaliação dos métodos de adição de detalhes - BD-PSNR

Tabela A.1: Avaliação das técnicas de adição de detalhes no método da malha triangular, com vista auxiliar H.

Sequência	BD-PSNR		%BR	
	Máscara	Direta	Máscara	Direta
<i>GT Fly</i>	0.912	0.897	-36.087	-35.237
<i>Poznan Hall 2</i>	0.419	0.155	-21.447	-10.901
<i>Pantomime</i>	0.709	0.558	-16.382	-13.628
<i>Champagne Tower</i>	0.566	0.555	-16.066	-15.723
<i>Balloons</i>	0.349	0.311	-8.949	-8.053
<i>Kendo</i>	0.307	-0.947	-7.338	27.860
<i>Book Arrival</i>	0.602	0.508	-17.708	-15.507
<i>Ballroom</i>	-0.016	-0.057	0.463	1.517
<i>Exit</i>	-0.163	-0.250	7.002	10.905

Tabela A.2: Avaliação das técnicas de adição de detalhes no método da malha triangular, com vista auxiliar V.

Sequência	BD-PSNR		%BR	
	Máscara	Direta	Máscara	Direta
<i>GT Fly</i>	1.560	1.725	-46.285	-46.733
<i>Poznan Hall 2</i>	0.043	-0.279	-2.151	23.862
<i>Pantomime</i>	0.065	-0.124	-1.559	3.194
<i>Champagne Tower</i>	0.242	0.219	-7.280	-6.658
<i>Balloons</i>	0.203	0.157	-4.890	-3.861
<i>Kendo</i>	0.199	-1.065	-4.917	33.958
<i>Book Arrival</i>	0.987	0.933	-26.426	-22.336
<i>Ballroom</i>	0.071	0.024	-1.814	-0.550
<i>Exit</i>	-0.064	-0.221	2.804	10.254

Tabela A.3: Avaliação das técnicas de adição de detalhes no método da pesquisa de blocos, com vista auxiliar H.

Sequência	BD-PSNR		%BR	
	Máscara	Direta	Máscara	Direta
<i>GT Fly</i>	0.769	0.762	-33.053	-32.929
<i>Poznan Hall 2</i>	0.480	0.438	-23.658	-21.986
<i>Pantomime</i>	-0.021	-0.029	0.677	0.916
<i>Champagne Tower</i>	0.390	0.386	-11.897	-11.793
<i>Balloons</i>	0.278	0.235	-7.342	-6.275
<i>Kendo</i>	0.269	0.220	-6.494	-5.405
<i>Book Arrival</i>	0.627	0.613	-18.319	-18.001
<i>Ballroom</i>	-0.011	-0.017	0.279	0.423
<i>Exit</i>	-0.008	-0.017	0.351	0.738

Tabela A.4: Avaliação das técnicas de adição de detalhes no método da pesquisa de blocos, com vista auxiliar V.

Sequência	BD-PSNR		%BR	
	Máscara	Direta	Máscara	Direta
<i>GT Fly</i>	1.503	1.589	-44.508	-45.622
<i>Poznan Hall 2</i>	0.006	-0.041	-0.130	2.843
<i>Pantomime</i>	-0.005	-0.007	0.117	0.191
<i>Champagne Tower</i>	0.055	0.053	-1.698	-1.617
<i>Balloons</i>	0.119	0.079	-2.953	-1.966
<i>Kendo</i>	0.138	0.085	-3.446	-2.167
<i>Book Arrival</i>	0.956	0.965	-25.809	-25.871
<i>Ballroom</i>	-0.012	-0.019	0.319	0.531
<i>Exit</i>	-0.004	-0.009	0.167	0.357

Tabela A.5: Avaliação das técnicas de adição de detalhes no método da pesquisa de blocos, com vista auxiliar H/V.

Sequência	BD-PSNR		%BR	
	Máscara	Direta	Máscara	Direta
<i>GT Fly</i>	1.378	1.487	-43.038	-43.386
<i>Poznan Hall 2</i>	0.447	0.399	-19.955	-18.125
<i>Pantomime</i>	-0.021	-0.028	0.602	0.814
<i>Champagne Tower</i>	0.364	0.361	-10.453	-10.347
<i>Balloons</i>	0.318	0.273	-7.576	-6.555
<i>Kendo</i>	0.313	0.261	-7.021	-5.948
<i>Book Arrival</i>	1.100	1.097	-28.438	-28.180
<i>Ballroom</i>	-0.014	-0.023	0.380	0.629
<i>Exit</i>	-0.008	-0.018	0.298	0.682

A.8 Avaliação dos métodos de melhoria da vista auxiliar

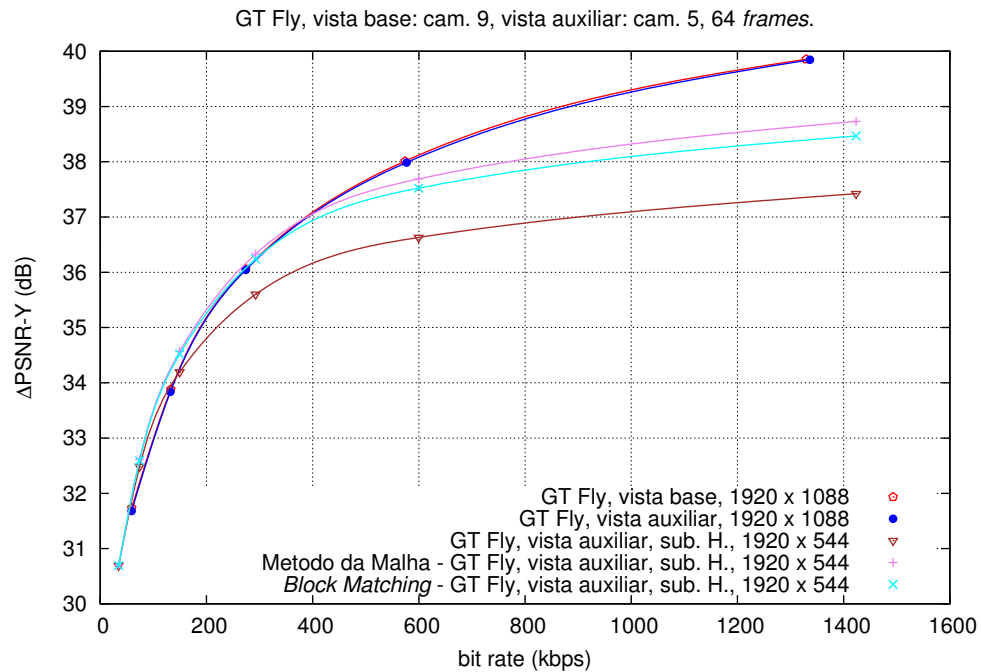


Figura A.133: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *GT Fly*, com vista auxiliar H.

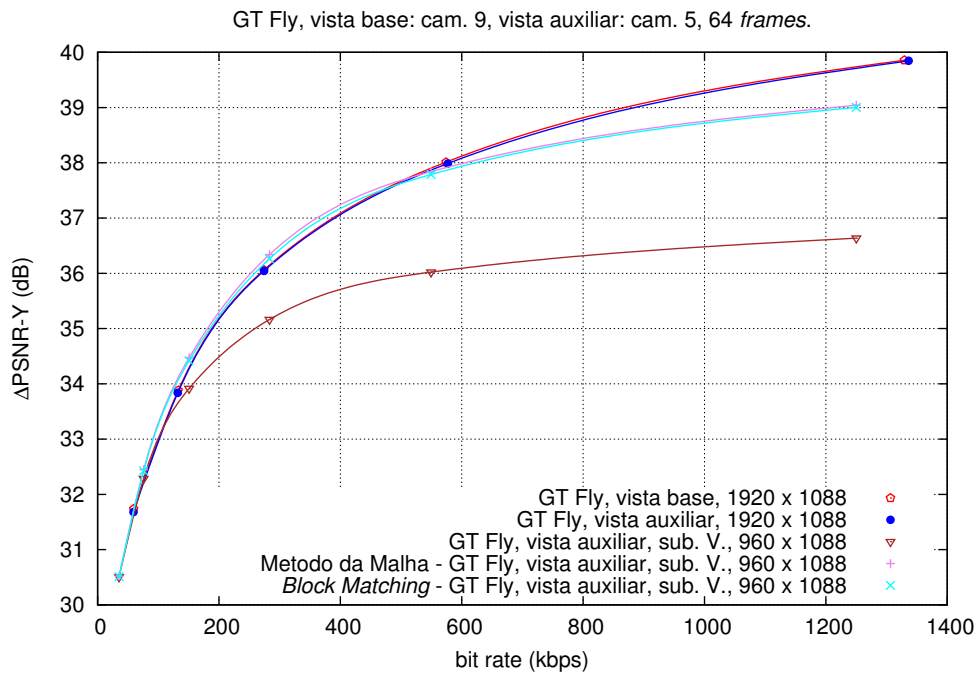


Figura A.134: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *GT Fly*, com vista auxiliar V.

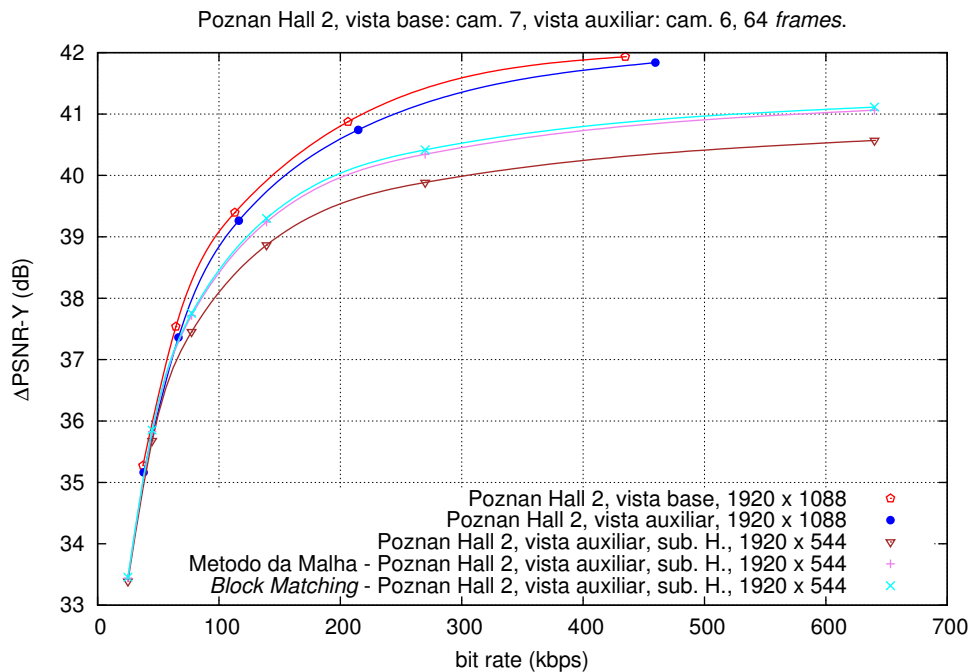


Figura A.135: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Poznan Hall 2*, com vista auxiliar H.

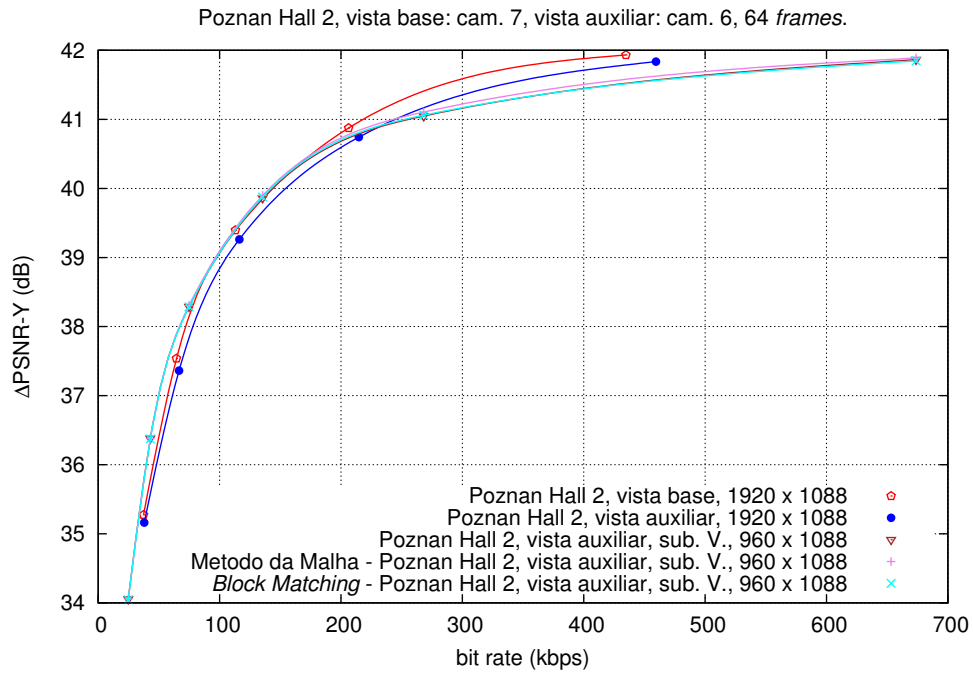


Figura A.136: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Poznan Hall 2*, com vista auxiliar V.

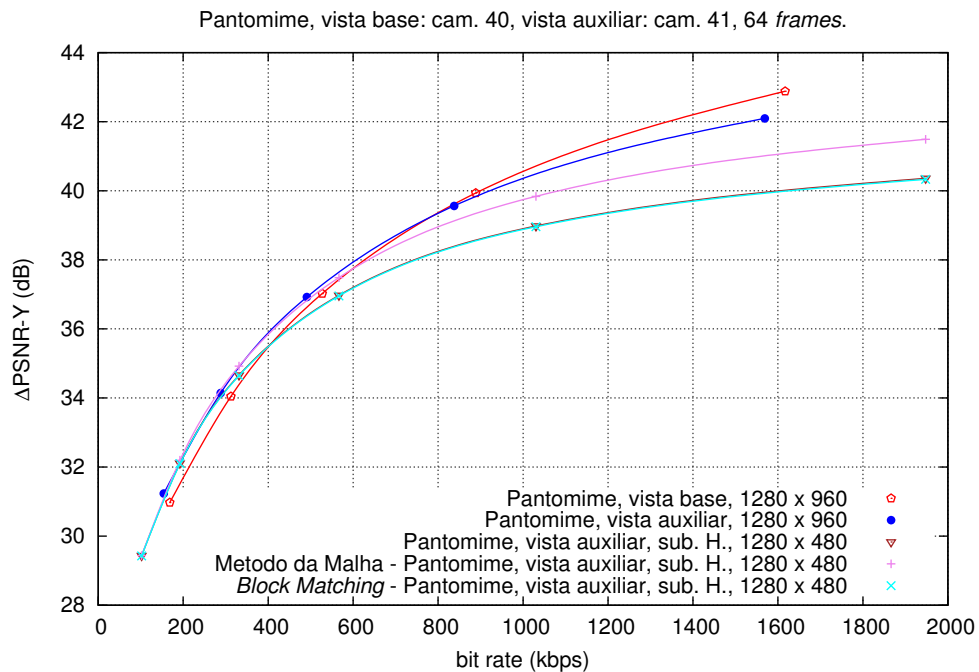


Figura A.137: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Pantomime*, com vista auxiliar H.

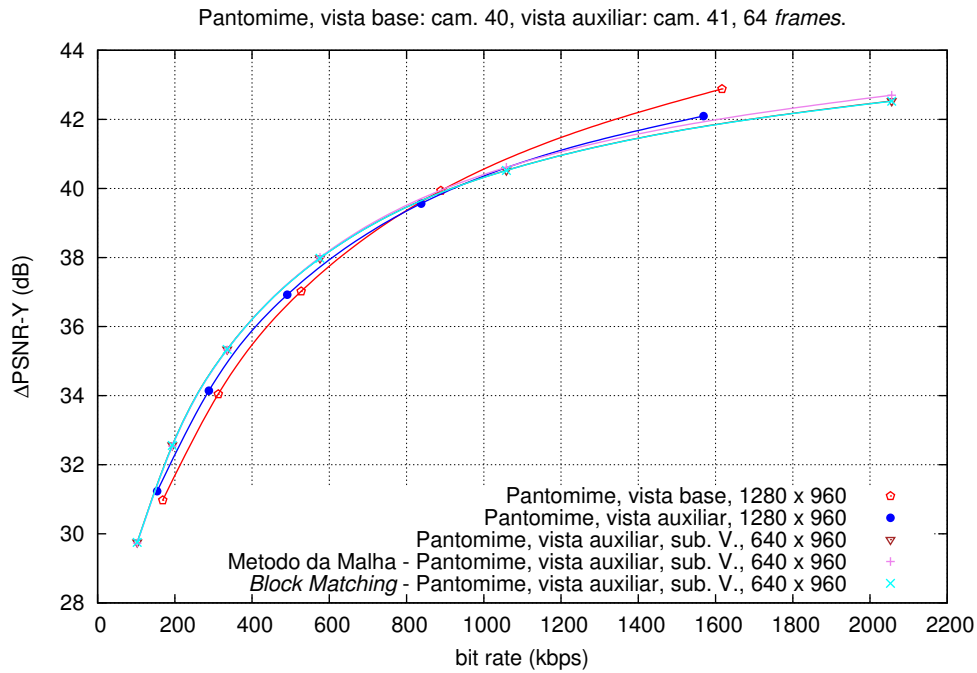


Figura A.138: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Pantomime*, com vista auxiliar V.

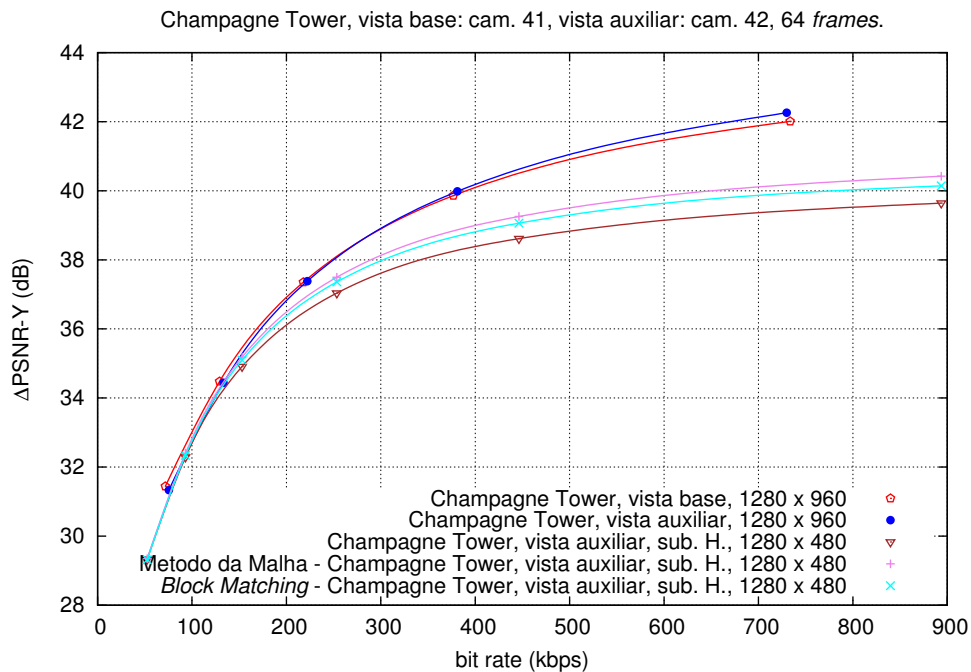


Figura A.139: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Champagne Tower*, com vista auxiliar H.

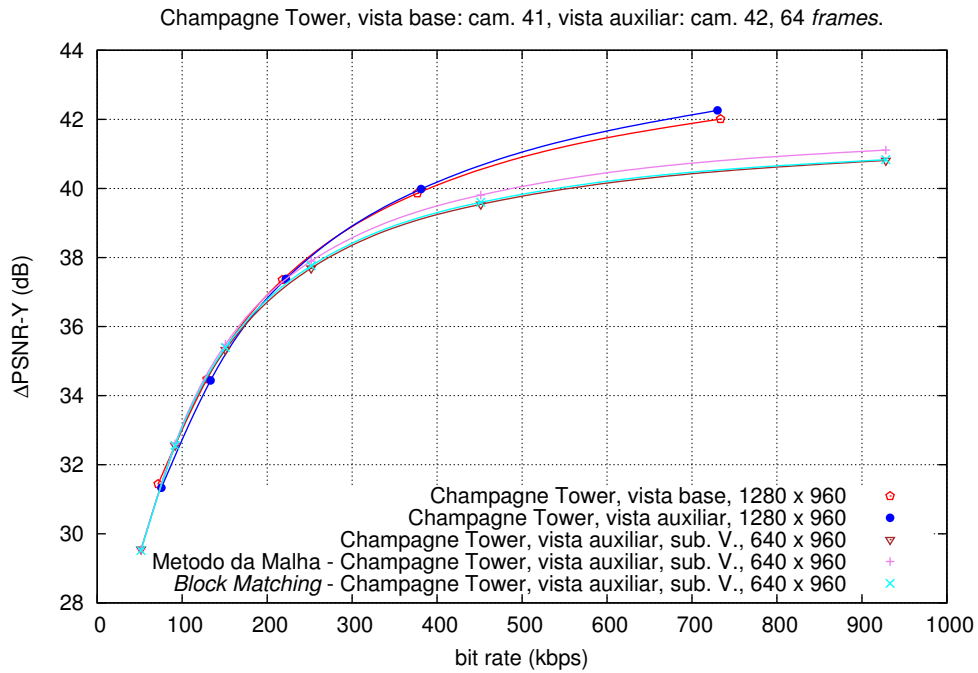


Figura A.140: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Champagne Tower*, com vista auxiliar V.

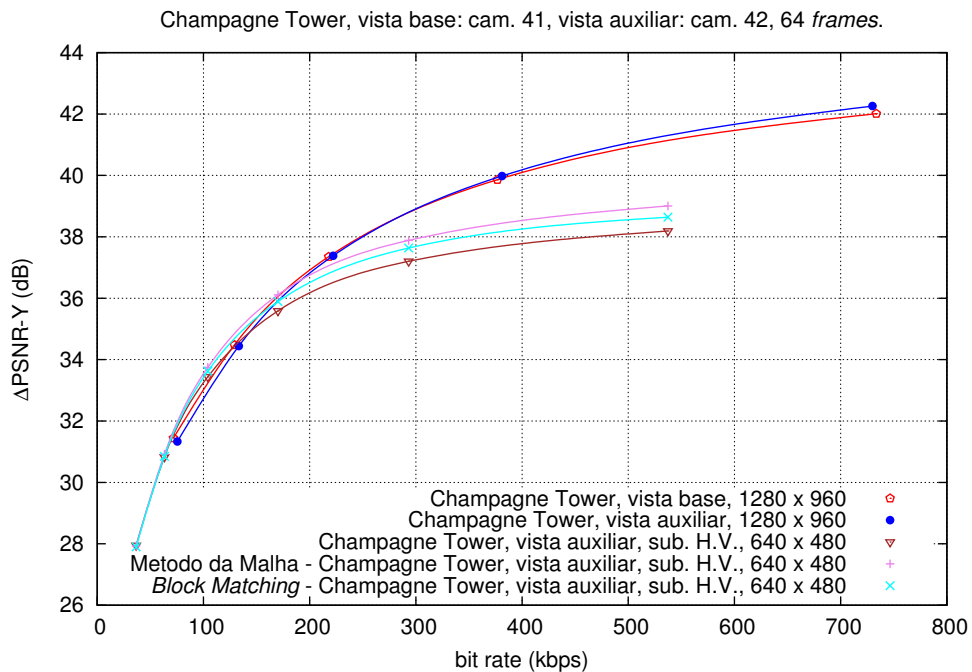


Figura A.141: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Champagne Tower*, com vista auxiliar H/V.

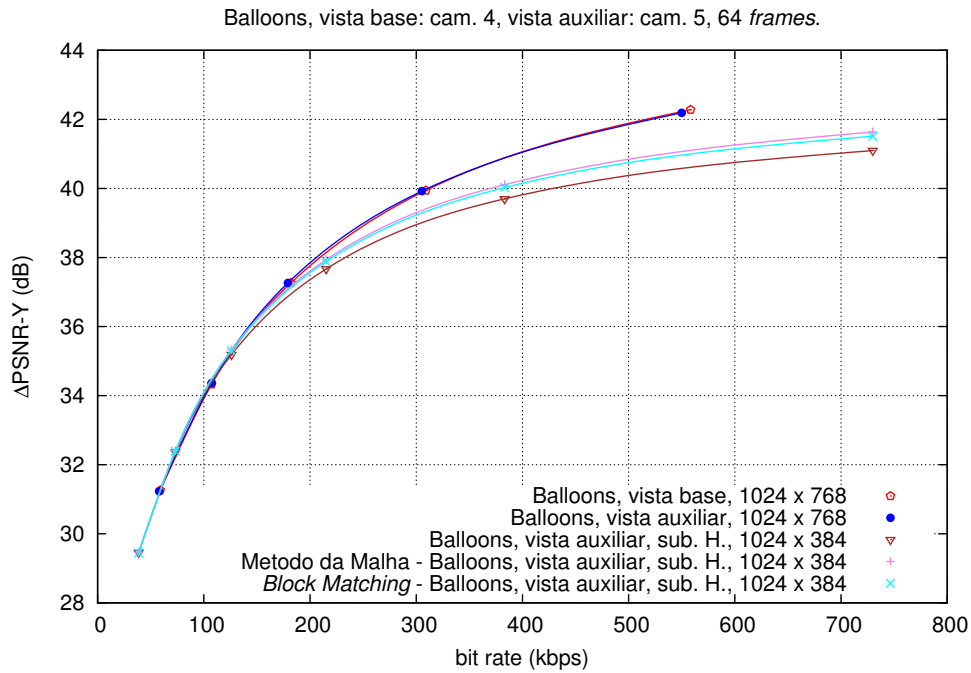


Figura A.142: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Balloons*, com vista auxiliar H.

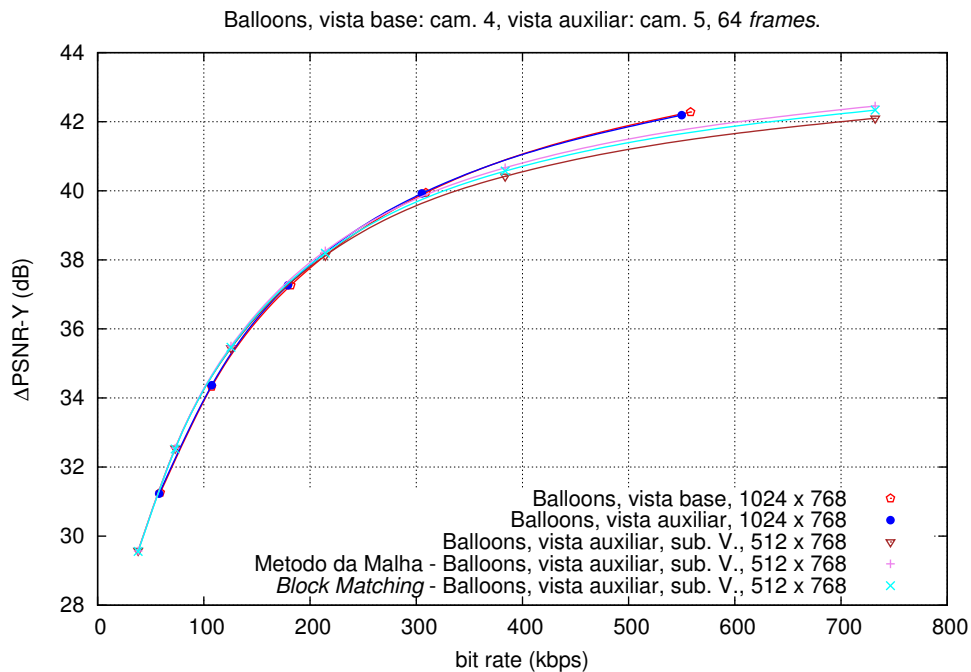


Figura A.143: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Balloons*, com vista auxiliar V.

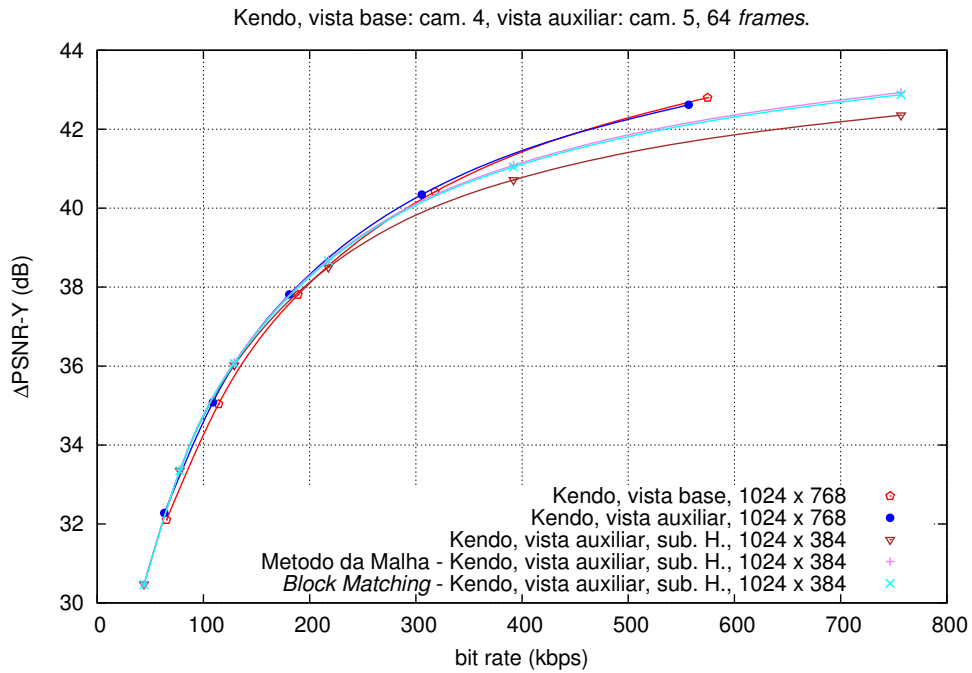


Figura A.144: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Kendo*, com vista auxiliar H.

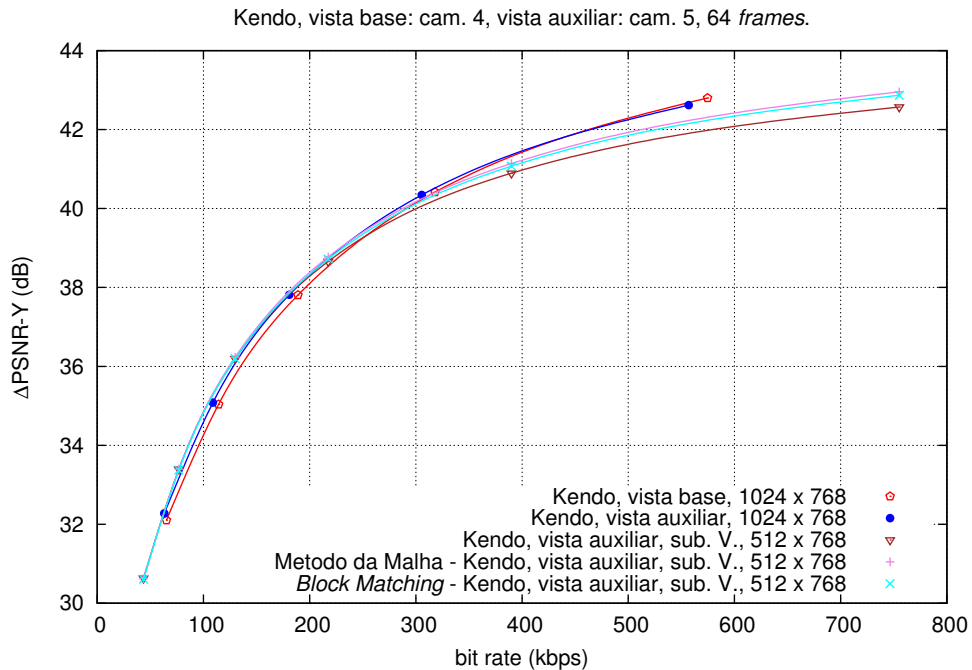


Figura A.145: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Kendo*, com vista auxiliar V.

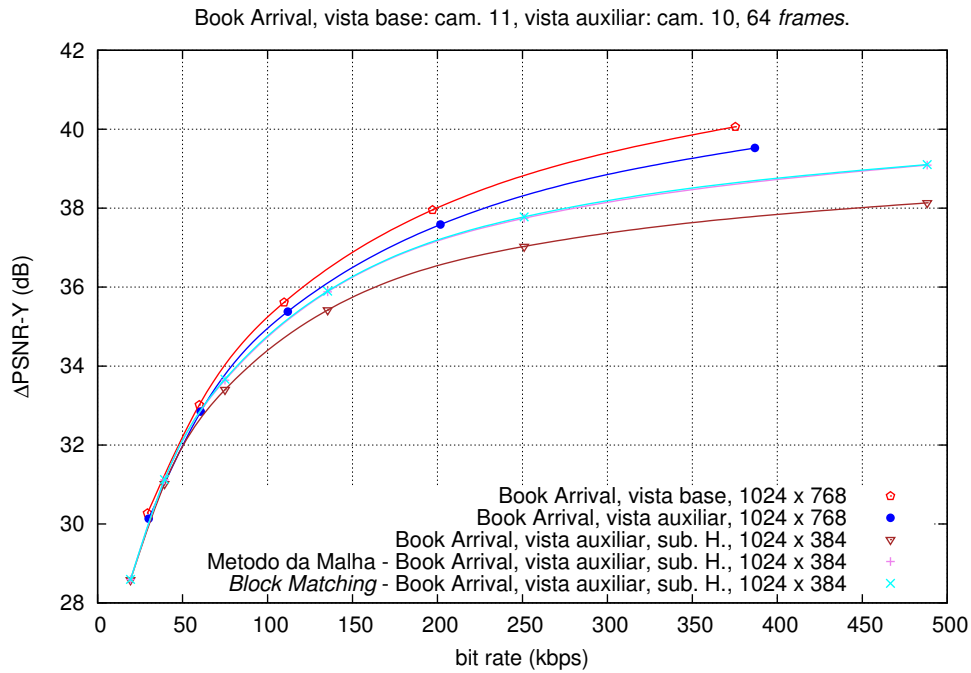


Figura A.146: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Book Arrival*, com vista auxiliar H.

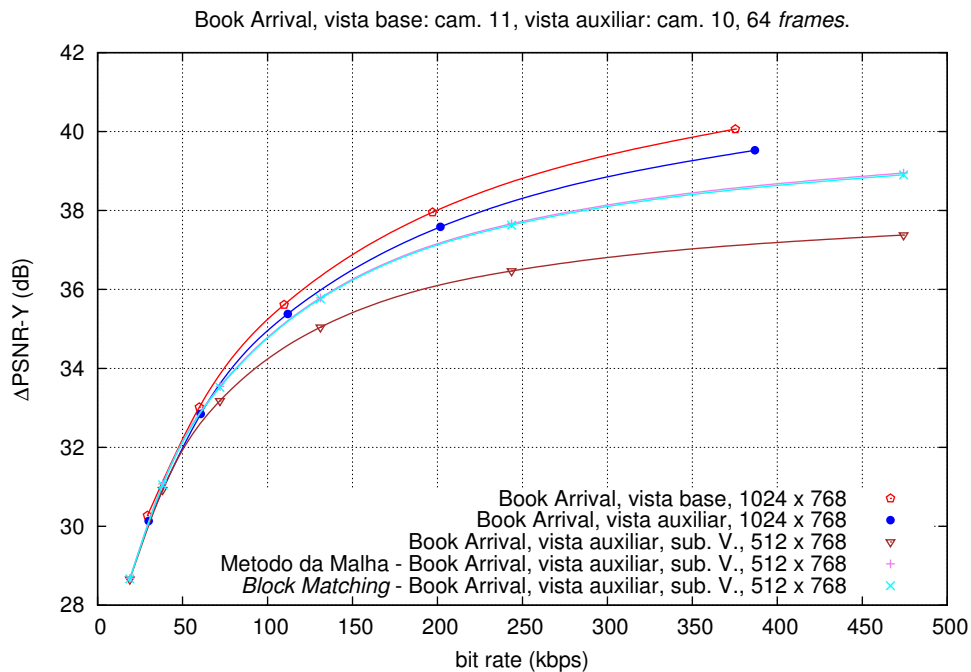


Figura A.147: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Book Arrival*, com vista auxiliar V.

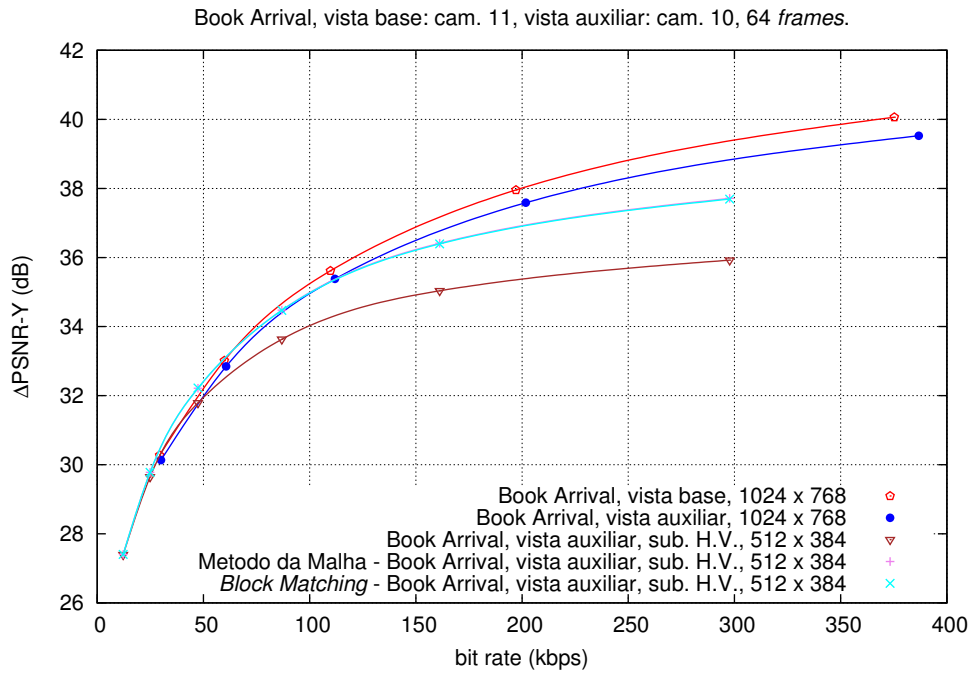


Figura A.148: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Book Arrival*, com vista auxiliar H/V.

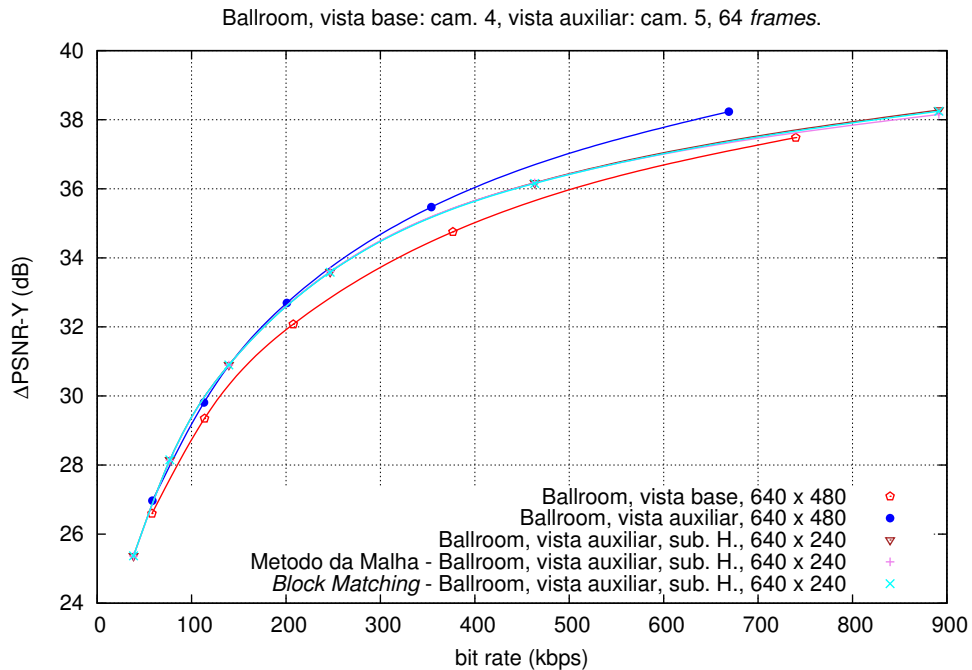


Figura A.149: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Ballroom*, com vista auxiliar H.

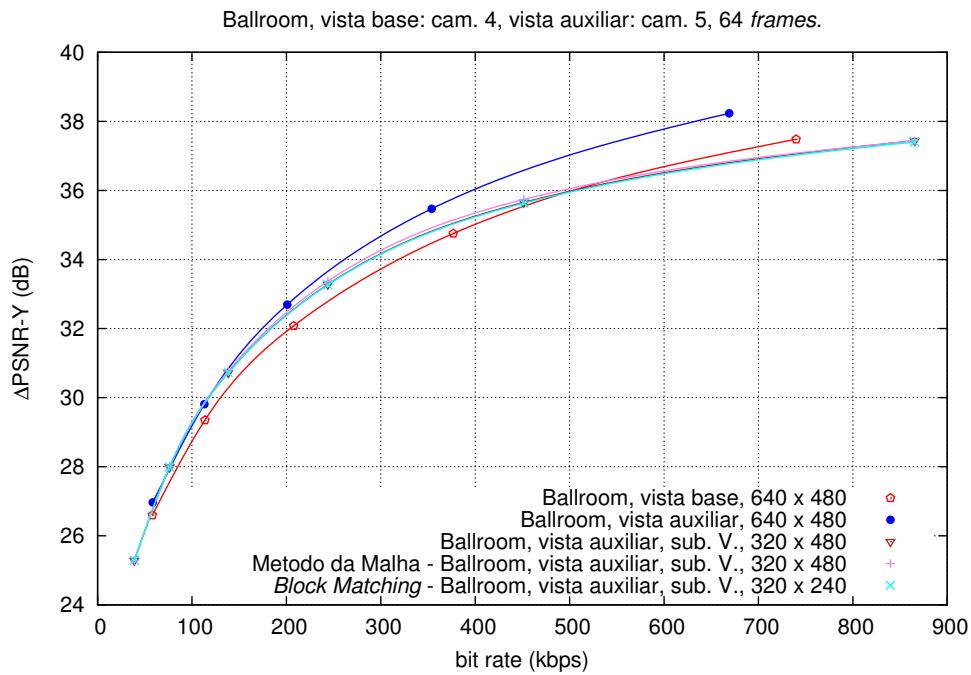


Figura A.150: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Ballroom*, com vista auxiliar V.

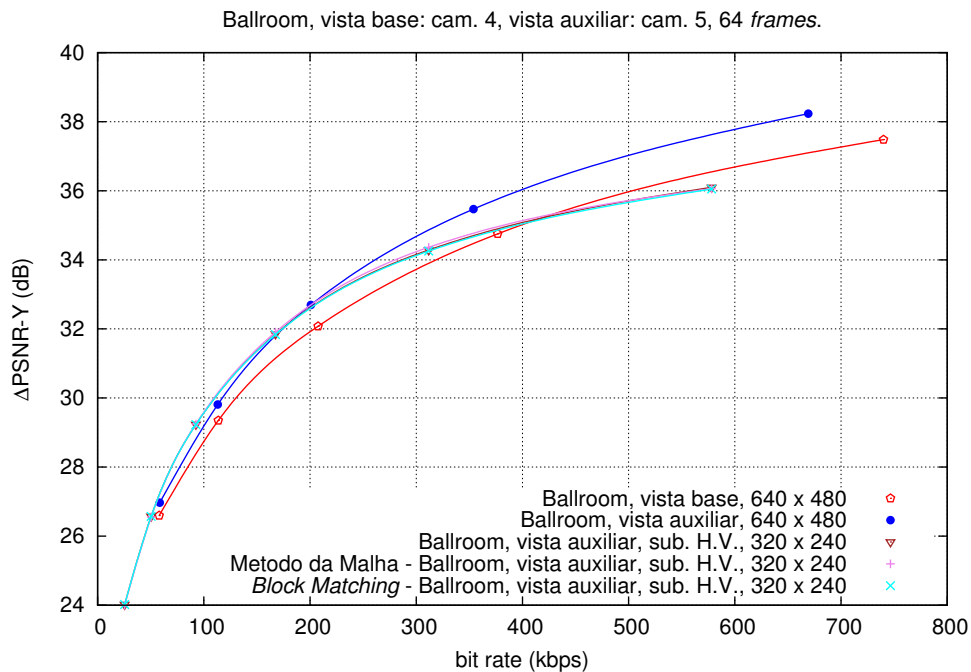


Figura A.151: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Ballroom*, com vista auxiliar H/V.

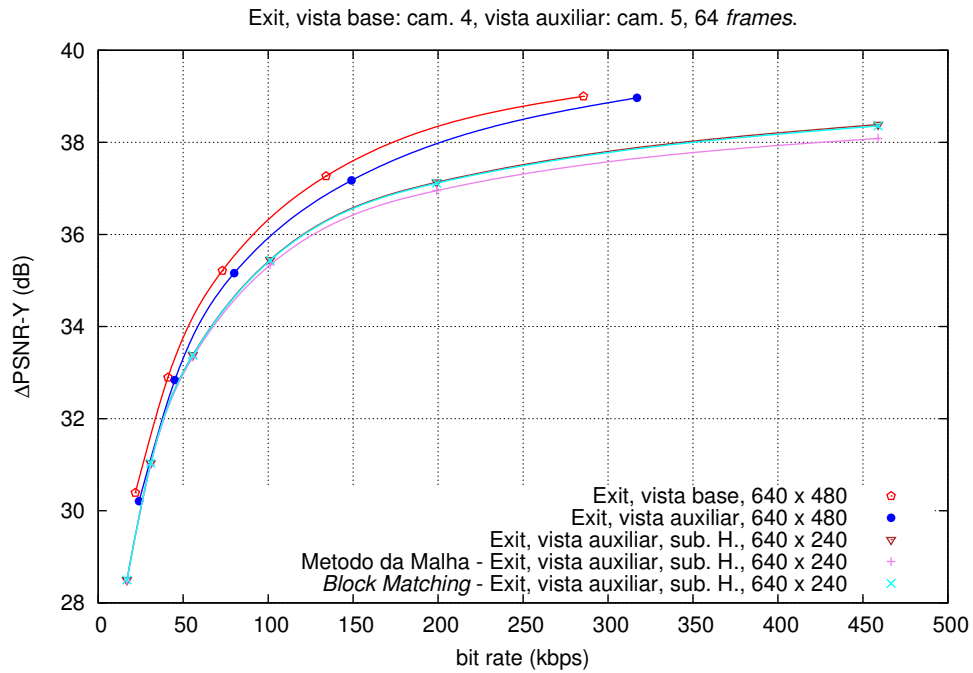


Figura A.152: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Exit*, com vista auxiliar H.

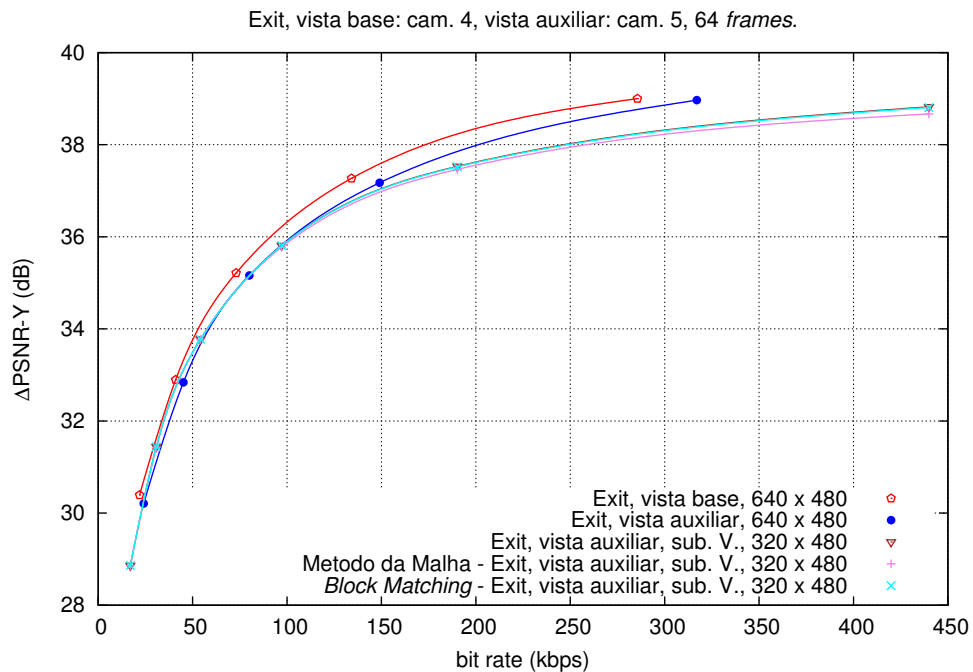


Figura A.153: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Exit*, com vista auxiliar V.

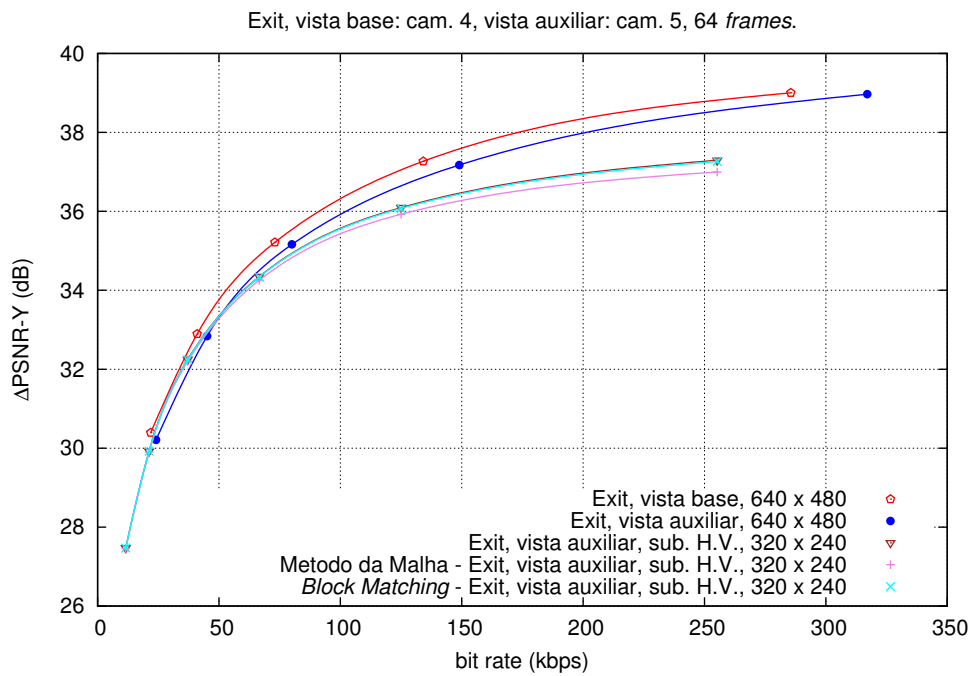


Figura A.154: Comparação entre o método baseado na malha e pesquisa de blocos, para a sequência *Exit*, com vista auxiliar H/V.

Apêndice B

Software de Simulação

B.1 Software de melhoria da vista auxiliar

Os programas realizados no decorrer do desenvolvimento do presente trabalho foram produzidos usando a linguagem C++ com o compilador *GNU Compiler Collection* (GCC) [40] em ambiente Linux. Foi utilizada a biblioteca de funções *Open Source Computer Vision Library* (OpenCV) [41], de modo a utilizar funções de processamento de imagem já implementadas, abreviando assim o tempo de desenvolvimento. O código desenvolvido não utiliza recursos específicos da plataforma utilizada, pelo que poderá ser utilizado em outro sistema operativo.

Foram desenvolvidos dois programas principais que implementam as técnicas em comparação nesta dissertação. O programa para efetuar a pesquisa de bloco semelhante assim como o programa para aplicar o método baseado na malha triangular, têm como dados de entrada as sequências de vídeo estéreo previamente codificadas e decodificadas com o HEVC. No caso da vista auxiliar foi aplicada a subamostragem antes da codificação e sobreamostragem após a decodificação. Após a codificação e decodificação, a sequência base foi sub/sobreamostrada. Deste modo os dados de entrada do programa correspondem aos 3 ficheiros de vídeo não comprimido (*raw*) YUV planar (4:2:0), relativos à vista base, vista auxiliar e vista base sub/sobreamostrada.

A subamostragem e sobreamostragem foi realizada através da ferramenta *DownConvertStatic* do software de referência *Joint Scalable Video Model* (JSVM) do *Scalable Video Coding* (SVC) (relativo à extensão anexo G do H.264/MPEG-4 AVC), disponível em [42]. A subamostragem e sobreamostragem foi realizada usando os parâmetros por omissão da ferramenta, ou seja, a subamostragem é baseada na função *Sinc Sine-windowed* e com a sobreamostragem é utilizado o método baseado num conjunto de filtros de 4 coeficientes inteiros derivados do filtro *Lanczos-3*.

A estrutura algorítmica dos programas segue a descrição dos métodos, já efetuada no Capítulo 4. De seguida efetua-se a descrição dos parâmetros disponíveis em cada programa, bem como a exemplificação de alguns testes.

O programa *Asymmetric Stereoscopic Video Enhancement – Mesh-based* (*asve_mb*) implementa o método baseado na malha triangular. O programa *Asymmetric Stereoscopic Video Enhancement – Block-based* (*asve_bb*) implementa o método baseado em pesquisa de blocos. Ambos os programas têm em comum a funcionalidade relativa aos seguintes parâmetros:

- `-i1 <infile1>`: nome do ficheiro de vídeo YUV com a vista base;
- `-i2 <infile2>`: nome do ficheiro de vídeo YUV com a vista auxiliar;
- `-i3 <infile3>`: nome do ficheiro de vídeo YUV com a vista base sub/sobreamostrada;
- `-r <orgrightview>`: ficheiro de vídeo com sequência auxiliar original para obtenção de PSNR;
- `-o <outfile>`: nome do ficheiro de saída para colocar a sequência auxiliar melhorada;
- `--help`: mostra um ecrã de ajuda na linha de comandos com a descrição dos parâmetros;
- `-w <width>`: largura da sequência em *pixels*;
- `-h <height>`: altura da sequência em *pixels*;
- `-s <searchWindowHor>`: tamanho da janela de pesquisa (largura) (valor por omissão: 40);
- `-v <searchWindowVert>`: tamanho da janela de pesquisa (altura) (valor por omissão: tamanho do bloco de pesquisa ou *kernel* + 4);
- `-q <ignoreHighFreq>`: nível de alta frequência da imagem de detalhes abaixo do qual que não deve ser adicionada (valor por omissão: 1);
- `-a <addDetailMethod>`: método de adição da imagem de detalhes (0: adição de toda a imagem de detalhes; 1: adição usando a imagem de contornos, 2: adição usando o PSNR) (valor por omissão: 1);
- `-f <framesToProcess>`: número de imagens da sequência de vídeo estéreo(*frames*) a processar.

Os seguintes parâmetros são usados apenas no programa *asve_mb*:

- **-k <dispKernSize>**: tamanho do *kernel* centrado nos vértices dos triângulos da malha de triangulação (valor por omissão: 11 (ou seja, 11x11));
- **-c <maxCorners>**: número máximo de cantos (*corners*) a gerar com o algoritmo *Shi-Tomasi* (valor por omissão: 5000).

No caso do programa *asve_mb* a linha de comando tem a seguinte forma geral:

```
$ ./asve_mb -i1 <infile1> -i2 <infile2> -i3 <infile3> -w <w> -h <h>
-o <outfile> [-p <processType> -r <orgRightView> -s <searchWindowHor>
-v <searchWindowVert> -k <dispKernSize> -c <maxCorners>
-q <ignoreHighFreq> -a <addDetailMethod> -f <framesToProcess>]
```

Exemplo de utilização para efetuar a melhoria da imagem auxiliar usando os parâmetros por omissão:

```
$ ./asve_mb -i1 balloons3_dec.yuv -i2 balloons5_decsub_up.yuv
-i3 balloons3_dec_subup.yuv -w 1024 -h 768 -o balloons5_dec_subup_mb_res.yuv
```

Os seguintes parâmetros são usados apenas no programa *asve_bb*:

- **-k <dbs_block_wdt>**: tamanho do bloco de pesquisa na horizontal (largura) (valor por omissão: 16);
- **-c <dbs_block_hgt>**: tamanho do bloco de pesquisa na vertical (altura) (valor por omissão: 8).

O programa *asve_bb* tem a seguinte forma geral para a linha de comando:

```
$ ./asve_bb -i1 <infile1> -i2 <infile2> -i3 <infile3> -w <w> -h <h>
-o <outfile> [-p <processType> -r <orgRightView> -s <search_window_hor>
-v <search_window_vert> -k <dbs_block_wdt> -c <dbs_block_hgt>
-q <ignoreHighFreq> -a <addDetailMethod> -f <framesToProcess>]
```

Exemplo de utilização para efetuar a melhoria da imagem auxiliar usando os parâmetros por omissão:

```
$ ./asve_bb -i1 balloons3_dec.yuv -i2 balloons5_decsub_up.yuv
-i3 balloons3_dec_subup.yuv -w 1024 -h 768 -o balloons5_dec_subup_bb_res.yuv
```

O código fonte do software desenvolvido, assim como os *scripts* utilizados, ficheiros de teste e resultados obtidos, encontram-se em anexo, armazenados em suporte ótico (CD).

Apêndice C

Publicações

C.1 Conferências

G. Jorge, S. M. M. Faria, N. M. M. Rodrigues, “Affine transform inter-view enhancement in spatial asymmetric stereo video decoding,” in *9th Conference on Telecommunications, Conftele 2013*, Castelo Branco, Portugal, May 2013.

Affine transform interview enhancement in spatial asymmetric stereo video decoding

Gilberto Jorge, Sérgio M. M. de Faria, Nuno M. M. Rodrigues
Instituto de Telecomunicações; ESTG, Instituto Politécnico de Leiria, Portugal
Email: gmj125@gmail.com, {nuno.rodrigues,sergio.faria}@ipleiria.pt

Abstract—In this paper we present a method to enhance, in the decoder, the reconstruction of a sub-sampled view in a HEVC-based simulcast transmission scenario of stereoscopic video sequences.

Despite maintaining compatibility with the legacy 2D video equipments, the simulcast transmission requires a high bit rate. In the absence of adequate 3D coding tools, it is possible to reduce the bit rate by using asymmetric video compression, where the base view is encoded at full spatial resolution and the auxiliary view is transmitted with a lower spatial resolution and then up-sampled at the decoder.

We have developed a method for enhancing the up-sampled auxiliary view at the decoder, that uses detail information from the base view (with full resolution). This process relies on Affine transforms to perform a geometric mapping between the high frequency detail in the full resolution base view and the low resolution auxiliary one. Additionally, in order to attain image continuity among regions and avoid blocking artifacts, the mapping uses a mesh triangulation from the auxiliary view applied to the detail image. The proposed technique is compared against the disparity block search method, and the results show that, for some sequences, it improves not only the objective quality (PSNR) by up to 1.7 dB, but also the subjective quality.

I. INTRODUCTION

The use of several views from the same scene to obtain 3D video leads to a significant increase in bandwidth requirements, when compared to monocular video. One way to reduce this bandwidth requirements is to use interview redundancy reduction techniques. Nevertheless, one may also exploit the features of the human visual system (HVS) for stereoscopic video by using asymmetric video compression, *i.e.*, the compression of the sequence's views with different quality levels. It has been shown that, if one of the views provides a sufficiently high quality level, the other view might be encoded with lower quality [1], since the brain will fuse those images in a way that still allows a convenient perception of depth [2]. To improve the stereoscopic acuity, the quality of the lower resolution view can be enhanced using post processing algorithms in the decoder.

The simulcast transmission of stereo video, based on asymmetric coding, that is coding one view with full spatial resolution and the other with half resolution (in horizontal and/or vertical direction), results in a bit stream reduction. After decoding and up-sampling the low resolution view, it presents a lower quality than the full resolution view, due to the lossy process occurred during sub-sampling, prior to coding.

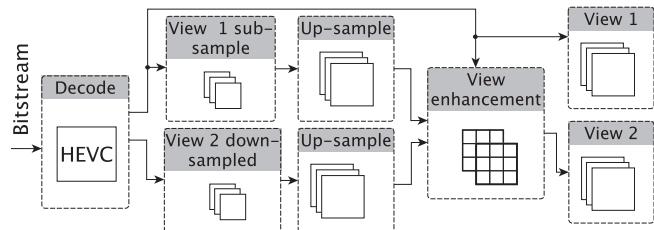


Fig. 1. General decoder architecture.

Some of the missing details may be present in the full resolution view, which can be used to enhance the auxiliary view. However this is not straightforward, since a spacial displacement (disparity) exists between a pair of views, which depends on the scene's depth, with respect to the cameras, and distance/angle between the cameras' optical axes. The conventional approach uses block-based methods to compensate this disparity.

The proposed method applies a triangular partition to the up-sampled low resolution image, and establishes a disparity mapping between triangles' vertices of both views. This operation is performed with Affine transforms, and the details contents are added from the high resolution image to the low resolution one. This method was compared with a block-matching based method that divides the image into fixed size blocks and uses the block disparity estimation to add details.

Both views, high and low resolution, have been encoded using the High Efficiency Video Coding (HEVC) [3] reference software (HM 9.1) [4].

II. PROPOSED METHOD

In the asymmetric simulcast transmission the base view is transmitted with full resolution; the auxiliary view is down-sampled prior to compression and up-sampled after being decoded, as can be seen in Fig. 1. In order to obtain the details (high frequencies) of the base view, the full resolution view is down-sampled and up-sampled after decoding. Subtracting this result to the full resolution view, the details image is obtained and can then be added to the up-sampled auxiliary view, *e.g.*, by using a block-matching method, or, as proposed in this paper, a triangular mesh mapping procedure.

A. Block-matching methods

In block-matching based enhancement, illustrated in Fig. 2, the auxiliary view is divided into fixed size square blocks.

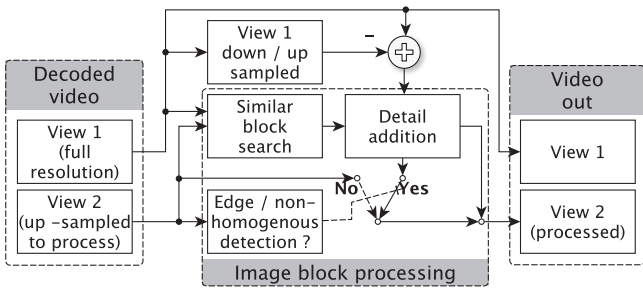


Fig. 2. Block-matching diagram.

Each block is used as a reference to search for a similar block in the base view, using disparity estimation, within a search window. This window is usually an horizontal rectangle, as the 3D stereo views sequences are typically acquired with horizontal disparity.

The disparity estimation is obtained comparing the smallest SAD (Sum of Absolute Differences) between the auxiliary view block and a block in the search window. The resulting value, for horizontal and vertical disparity of each block, is used as offset for the displacement of the block in the full resolution detail's image block. This block can then be added to the decoded up-sampled auxiliary view. The decision of adding a detail block depends on a mask that signalizes the (usually non-homogeneous) areas which may benefit from the added details. The use of this mask is explained in a future section.

B. Mesh-based method

In this paper we propose a sub-sampled view enhancement method based on the mapping of a triangulation mesh, created over the base view image edges. By using the warping of a triangular mesh, the proposed algorithm reduces the block effect of the previous method, thus improving the reconstructed image quality. First is applied the Canny edge detection algorithm [5] to the auxiliary view image and then the Shi-Tomasi [6] corner detection algorithm to obtain the image corner points, as represented in Fig. 3. Afterwards, the Delaunay triangulation is applied to the auxiliary view image using the previously determined set of corner points. This operation results in a non-regular tessellation with triangles which limit different texture areas, as can be seen in Fig. 4.

The non-regular tessellation of the full resolution view is formed by displacing the auxiliary view's triangles' vertices, according to the disparity between both views. The disparity estimation is determined by comparing the smallest SAD between the block centered in each vertex, in the auxiliary view, and the matching vertex kernel block in a search window, defined in the base view.

Finally, an Affine transform is applied using the triangles' vertices coordinates and the respective disparity, on the details image. The resulting image with details from the full resolution view displaced and transformed is added to the up-sampled low resolution view.

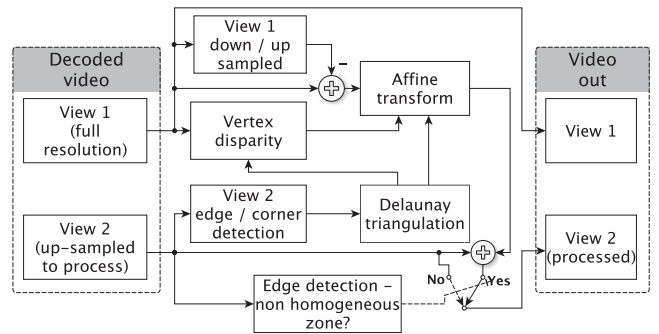


Fig. 3. Proposed post-processing method.

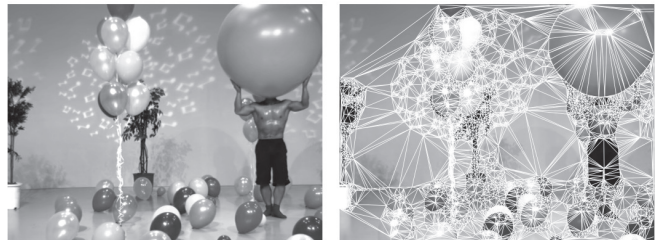


Fig. 4. Delaunay triangulation applied to image 1 of Balloons sequence.

C. Details image addition

Mesh-based and Block-matching methods use the same details image addition process. Due to the differences between the base and auxiliary views, the addition of the base view details is not always beneficial for the auxiliary view. In order to evaluate which regions of the full resolution image contribute to enhance the quality of the up-sampled auxiliary image, an auxiliary test was conducted, using the original auxiliary view image. In this procedure, the image was partitioned into fixed size square blocks, that are individually processed. The details from the high resolution image were only used when the auxiliary view's block reconstructed with the high resolution image details achieved higher objective quality (PSNR) than the block reconstructed with the low resolution image alone. For example, the up-sampled image represented on the left in Fig. 5 accepts 30% of its blocks to be enhanced (white) with details from the high resolution view (image on the right).

As can be observed in this example, most chosen blocks lie on non-homogeneous areas, e.g., in the objects' boundaries. But in these areas it also may be difficult to recover information, as these regions are more subject to occlusions due to the disparity. We can also note that there are some details in homogeneous areas that are worthy to be added to the auxiliary view, as illustrated in the picture. But, as expected, when the triangle vertices occur in flat areas (or blocks in Block-matching), it is more difficult to estimate the disparity, due to the homogeneity of the pixel values.

In a practical scenario, the original auxiliary view is obviously not available in the decoder, so we have developed an algorithm to decide which regions should be chosen to add-up details from the high resolution view. Firstly, an edge detection



Fig. 5. Mode decision using the original image.

algorithm is applied to the decoded up-sampled view. Then, dividing the same image view in blocks and only adding the details if the corresponding given block contains an edge, a mask is created around these regions, in order to allow the details to be added only in the selected edge areas.

III. EXPERIMENTAL RESULTS

Experimental tests were carried out using publicly available video test sequences, encoded with the HEVC test model (HM 9.1). 64 frames of the used sequences were encoded using the random access high efficiency configuration (10 bit), with the quantization parameters (QP) used in the HEVC core experiments common test conditions [7]: 22, 27, 32 and 37.

Prior to the evaluation of the rate-distortion performance of the proposed methods, the configuration parameters for the proposed methods were established.

A. Mesh-based configuration

The kernel size used to perform the disparity estimation of each vertex of the triangle is very important, as it will influence the operation accuracy. The search window size is another relevant parameter. Since the used video sequences were obtained with an aligned camera arrangement, the search window height was determined by the search block size plus two pixels of tolerance in the top and bottom, respectively.

Initial tests indicated that the best results are obtained using kernels with sizes 7x7, 9x9, 11x11 and 13x13, so a first test was conducted in order to determine which of these values was the best. Table I shows the difference in dB, for various QPs, between the processed low quality sequence and the non-processed one, for two stereo sequences, Kendo and Balloons (cameras 4 - full resolution, and 5 - down-sampled view, for both stereo sequences). The auxiliary view was down-sampled by a factor of two in both horizontal and vertical. Tests were performed using additional sequences (BookArrival, Pantomime, Ballroom), with similar results. The kernel size was tested for the first five frames of each sequence. Initially, a search window of 40 by (kernel size+4) pixels for width and height, respectively, was used. The conducted tests indicated that the overall best vertex kernel size was 11x11.

Using the vertices kernel size of 11x11, the next step was to determine the search window width. Various sizes of search windows were tested, from 18 to 48, using QP=22 and the horizontal down-sampling direction. These tests indicate that this parameter is not relevant, as long as its value exceeds

TABLE I
VERTEX KERNEL FOR DISPARITY ESTIMATION

Seq.	Kernel	QP=22	QP=27	QP=32	QP=37
Kendo cam. 4, 5	7	0.63	0.42	0.19	0.06
	9	0.70	0.48	0.23	0.08
	11	0.78	0.53	0.27	0.09
	13	0.82	0.54	0.28	0.11
Balloons cam. 4, 5	7	0.63	0.48	0.25	0.10
	9	0.66	0.50	0.30	0.13
	11	0.67	0.50	0.31	0.13
	13	0.63	0.48	0.30	0.14

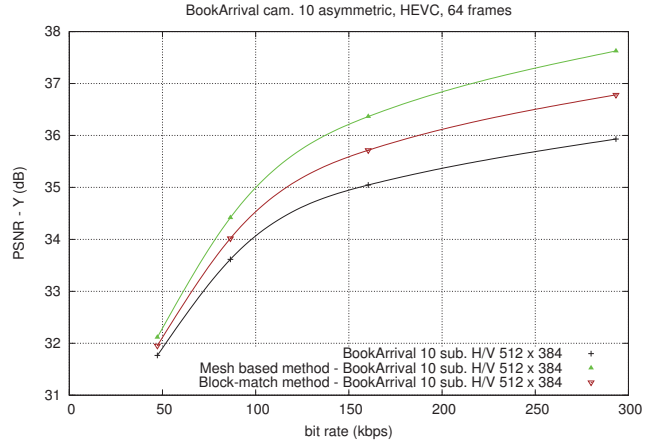


Fig. 6. Book Arrival sequence using Mesh and Block-matching method.

the stereo disparity. In this case the search window, for next comparison tests with Mesh method, has been defined with the value 40. However, a larger value may be set for a pair of views taken from cameras with a larger optical axes.

B. Block-matching parameters

The parameters for the block-matching based method were obtained in a similar way as for the triangulation method. In this case, for each search window, various block sizes were tested. The 16x8 block revealed the higher gain for all down-sampled directions. The chosen search window was 40 pixels of width, and the size of the block plus two pixels in top and bottom for the height.

C. Results Discussion

The rate-distortion results are presented for the stereo sequence Book Arrival, in Fig. 6. From this image and the results in Table II, we can observe that both enhancement methods increase the quality of the up-sampled image and that the Mesh method consistently achieves better results than the Block-matching method, according to the Bjontegaard Delta PSNR (BD-PSNR) and percentage of bit rate gain (%BR)[8].

Tests conducted for auxiliary view used down-sampling in the horizontal (H) and vertical (V) directions, as well as in both of the H and V (H/V). The results showed that the H/V method attained higher enhancement than the other down-sampling methods (H and V) alone. This also reflects

TABLE II
EVALUATION OF ENHANCEMENT METHODS.

Sequence (resolution)	DB-PSNR Mesh	DB-PSNR Block	%BR Mesh	%BR Block
BookArrival (1024x768)	1.051	0.532	-26.91	-17.01
Pantomime (1280x960)	0.555	-0.005	-12.31	0.15
Kendo (1024x768)	0.377	0.306	-8.22	-6.80
Balloons (1024x768)	0.414	0.330	-9.32	-7.64
Ballroom (640x480)	0.079	-0.228	-1.97	5.96
Exit (640x480)	-0.126	-0.382	4.79	14.30

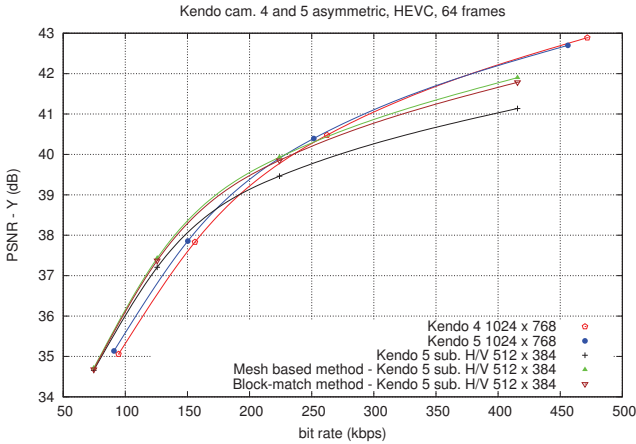


Fig. 7. Kendo sequence using Mesh and Block-matching method.

that higher rate gains may be achieved when there is more information discarded (H/V). The results presented in this section used a H/V sub-sampling. In all sequences tested, the Mesh-based method achieves better objective quality than the Block-matching method, that is 0.37 dB for Kendo, and 1.05 dB for Book Arrival sequence, as can be seen in Table II.

Another aspect that deserves to be discussed is the dependence on the images' content and spatial resolution. For sequences with lower spatial resolution, the quality gains of image enhancement are lower, as the image content already carries less high frequency details, thus the high resolution image is useless. As can be observed from the rate-distortion performance evaluation both methods may have a clearly different behavior, depending on the sequence contents.

As can be observed in Fig. 7, for some sequences there is bit rate value below which it is worthwhile to use asymmetric view coding, as lower bit rates can be achieved in the aggregate of both views, assuming that the auxiliary view may be encoded with lower quality. If a value of objective image quality in terms of PSNR can be established for such quality difference between both views, then the bit rate threshold can be determined for asymmetric view coding.

Regarding the subjective evaluation of both methods, we can observe in Fig. 8 a sample of image enhancement, comparing a region of image 1 of Book Arrival sequence (cam. 10), where the processed image was down-sampled in H/V. While the down-sampled version appears blurred, the sample processed by the Block-based presents the details misplaced causing

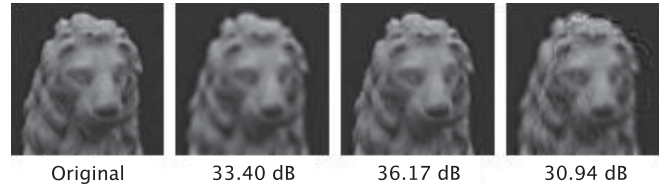


Fig. 8. Subjective and objective comparison on Book Arrival image detail (cam. 10, frame 1) (left to right: Original, Up-sampled only, Mesh-based, Block-based).

artifacts, unlike Mesh-based one.

IV. CONCLUSION

This paper discusses the use of a Mesh-based method to enhance the quality of an up-sampled view in an asymmetric stereoscopic video transmission. The proposed method aims to encode a pair of views reducing the bit rate required to transmit a stereo sequence. It improves most stereo sequences asymmetrically encoded, with different resolutions, except low resolution ones and those with significant content occlusions between both views.

Comparing the developed method with a simple disparity Block-matching based method, gains of up to 1.05 dB were achieved, in terms of Bjontegaard Delta PSNR.

Future developments include experimental tests to verify possible differences in depth perception of the improved sequences, through subjective testing.

ACKNOWLEDGMENT

This project was funded by Instituto de Telecomunicações (IT), in the ambit of the project PTDC/EEA-TEL/099387/2008 with the title "COMUVI-Compressão de vídeo multi-vista".

REFERENCES

- [1] L. B. Stelmach, W. J. Tam, D. V. Meegan, A. Vincent, and P. Coriveau, "Human perception of mismatched stereoscopic 3d inputs," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2000, pp. 5–8.
- [2] M. G. Perkins, "Data Compression of Stereopairs," in *IEEE Transaction on Communications*, 40, pp. 684–696, 1992.
- [3] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," in *IEEE Transactions on Circuits and Systems for Video Technology*, May 2012.
- [4] I. K. Kim, K. McCann, K. Sugimoto, B. Bros, and W. J. Han, "Hm7: High efficiency video coding (HEVC) test model 8 encoder description," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 10th Meeting*, Stockholm, SE, Jul 2012.
- [5] J. Canny, "A computational approach to edge detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-8, Issue 6*, November 1986, pp. 679 – 698.
- [6] J. Shi and C. Tomasi, "Good features to track," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, Seattle, Washington, June 1994, pp. 593 – 600.
- [7] F. Bossen, "Common test conditions and software reference configurations," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 10th Meeting*, Stockholm, SE, Jul 2012.
- [8] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *VCEG Contribution VCEG-M33*, Austin, April 2001.