University of Texas Rio Grande Valley ScholarWorks @ UTRGV

Theses and Dissertations - UTB/UTPA

8-2005

Entropy-based 2D image dissimilarity measure

Meng-Hung Wu University of Texas-Pan American

Follow this and additional works at: https://scholarworks.utrgv.edu/leg_etd

Part of the Computer Sciences Commons

Recommended Citation

Wu, Meng-Hung, "Entropy-based 2D image dissimilarity measure" (2005). *Theses and Dissertations - UTB/ UTPA*. 777.

https://scholarworks.utrgv.edu/leg_etd/777

This Thesis is brought to you for free and open access by ScholarWorks @ UTRGV. It has been accepted for inclusion in Theses and Dissertations - UTB/UTPA by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact justin.white@utrgv.edu, william.flores01@utrgv.edu.

ENTROPY-BASED 2D IMAGE DISSIMILARITY MEASURE

A Thesis

by

MENG-HUNG WU

Submitted to the Graduate School of the University of Texas – Pan American In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2005

Major Subject: Computer Science

ENTROPY-BASED 2D IMAGE DISSIMILARITY MEASURE

A Thesis by MENG-HUNG WU

Approved as to style and content by:

Zaci, Jox Ping-Sing Tsai, Ph. D. Chair of Committee

Zhixiang Chen, Ph. D. Committee Member

Peter A. Ng, Ph. D. Committee Member

August 2005

ABSTRACT

Wu, Meng-Hung, <u>Entropy-Based 2D Image Dissimilarity Measure</u>. Master of Science (MS), August, 2005, 57 pp., 5 Tables, 20 Figures, references, 12 titles.

Traditional histogram or statistics based 2D image similarity/dissimilarity metrics fail to handle conjugate pair of black and white images, due to the lack of spatial information in the measurement. Recently proposed Compression-based Dissimilarity Measure (CDM) [1] based on the concept of Kolmogorov complexity has provided a different paradise for similarity measurement. However, without a clear definition how to "concatenate" two 2D images, CDM has difficulties to directly apply with 2D images. In this thesis, an *entropy*-based 2D image dissimilarity measure is proposed within the same Kolmogorov complexity paradise. The spatial relationship between images is embedded in our metric, and the actual compression of images is not needed once the entropy values are obtained. The proposed metric has been tested for scene change detection application, and encouraging results are presented here.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Zhixiang Chen of the department of Computer Science UTPA for his useful discussion, Mr. David Kirtley for his help of recoding the test video sequences, and Mr. Jason Sada of GridIron Technologies LLC, Scottsdale, AZ for providing test images.

TABLE OF CONTENTS

Abstractiii
Acknowledgmentsiv
Table of Contents
List of tableS
List of Figures
Chapter I 1
Introduction1
Chapter II 3
Background and Related Works 3
Chapter III
Entropy-BASED 2D Image Dissimilarity Measure
Chapter IV 10
Experimental Data and Results10
Still Frame Comparisons 11
Video Scene Change Detection
Video Sequence 1 15
Video Sequence 2
Video Sequence 3
Chapter V

·....

Conclusions	45
References	46
Vita	48

LIST OF TABLES

Table 1 Summary of Test Metrics	7
Table 2 Still Images Comparison	14
Table 3 Dissimilarity Plot for Each Metric for Video Sequence 1	19
Table 4 Dissimilarity Plot for Each Metric for Video Sequence 2	30
Table 5 Dissimilarity Plot for Each Metric for Video Sequence 3	43

LIST OF FIGURES

Figure 1 Conjugate pair of black and white images
Figure 2 Still images
Figure 3 Frames 16 to 18 from video sequence 1
Figure 4 Frames 31 to 34 from video sequence 1 17
Figure 5 Frames 42 to 44 from video sequence 1
Figure 6 Frames 9 to 20 from video sequence 2
Figure 7 Frames 50 to 57 from video sequence 2
Figure 8 Frames 60 to 63 from video sequence 2
Figure 9 Frames 156 to 159 from video sequence 2
Figure 10 Frames 205 to 207 from video sequence 2
Figure 11 Frames 288 to 297 from video sequence 2
Figure 12 Frames350 to 355 from video sequence 2
Figure 13 Frames 23 to 28 from video sequence 3
Figure 14 Frames 32 to 37 from video sequence 3
Figure 15 Frames 59 to 66 from video sequence 3
Figure 16 Frames 97 to 110 from video sequence 3
Figure 17 Frames 148 to 155 from video sequence 3
Figure 18 Frames 181 to 188 from video sequence 3
Figure 19 Frames 203 to 214 from video sequence 3

~

CHAPTER I

INTRODUCTION

Similarity/Dissimilarity measures are required in many different applications, such as data mining, time series classification/anomaly detection [2], motion estimation in video coding, scene change detection, and so on. In general, when applied to 2D images, these measures fall into four different categories: histogram based methods, first and/or second order statistic based methods, template based methods, and compression based methods.

The traditional histogram [3, 4] or statistic based metrics [5, 6, 7] fail to handle conjugate pair of black and white images (as shown in Figure 1) due to the lack of 2D spatial information in the measurement. Template based metric, as discussed in [4], usually requires high computational complexity and has huge ranges for the measure value. Typical compression based metric depends upon the underline compression methods. For example, as in [8], the DC coefficients of the discrete cosine transform (DCT) from the MPEG stream are used for video scene change detection. It will be difficult to extend the same metric to video stream with a discrete wavelet transform (DWT) based video compression such as motion JPEG2000 [9]. The newly proposed compression-based dissimilarity measure (CDM) provides a different metric based on the concept of Kolmogorov complexity. CDM has been tested extensively on time series data

Amonda .

for clustering. However, without a clear definition on how to "concatenate" two 2D images, CDM has difficulties to directly apply with 2D images for similarity/dissimilarity measure.

In this thesis, we propose an *entropy*-based 2D image dissimilarity metric inspired by the CDM. The spatial relationship between images is embedded in our metric, and we do not need to compress the image once the entropy values are obtained. The rest of this thesis is organized as follows. Chapter II presents the background and related works on CDM, entropy, and video scene change detection. In Chapter III, we present the proposed new entropy-based 2D image dissimilarity metric. Experimental results are presented in Chapter IV, and a conclusion is provided in Chapter V.

CHAPTER II

BACKGROUND AND RELATED WORKS

Before proposing our new 2D images dissimilarity metric, we would review the two key major components, *compression-based dissimilarity measure* and *entropy*, which are the building blocks for the proposed measure. We also review the basic metrics used for video scene change detection.

1. Compression-based Dissimilarity Measure (CDM)

In [1], based on the concept of Kolmogorov complexity, the authors present the CDM metric as follows

$$CDM(x, y) = \frac{C(xy)}{C(x) - C(y)}$$

Where x and y are two strings, xy is the concatenation of the two strings, and $C(\bullet)$ is the compressed size of the input data. Intensive experiments with successful results were reported. However, when applying the CDM with 2D images, the concatenation operation is not clearly defined even when the two images are the same size. Also the text compressors (zip, gzip, compress, bzip2, etc.) mentioned in [1] are either statistical or dictionary based methods. Any image compressor (lossless) will try to

reduce not just the statistical redundancy but also the spatial redundancy. De-correlation type transformation may also be applied for image compression. All these make the CDM difficult apply to 2D images directly.

2. Entropy

The term "entropy", as defined in Merriam-Webster dictionary, means: a measure of the unavailable energy in a closed thermodynamic system that is also usually considered to be a measure of the system's disorder or chaos. However, from the Information Theory [10] point of view, *entropy* is the expected length of a binary code over all possible symbols in a discrete memory-less source. In other words, entropy can be considered as the average number of bits one needs to represent a symbol in a stationary system, where the limited source symbols have fixed probabilities of occurrence. The entropy is expressed as

$$E = -\sum_{i=1}^{N} p(a_i) \log_2 p(a_i)$$

Where N is number of symbols and $p(a_i)$ is the probability of occurrence of symbol a_i . This is a very convenient measure for any coding system, and it provides a bound for compression that can be achieved.

The entropy of an image can be easily calculated based on the image histogram information, which is nothing but the occurrence information of all the intensity values

(symbols) in the image. Instead of compressed size of the input data used in CDM, we will use entropy in our proposed metric.

3. Video Scene Change Detection

In order to browse through videos or store them into indexed databases, image sequences must be segmented or indexed into smaller clips through the technique of scene change detection. Scene changes are transitions in the video field contents and can occur in the transition of a single frame (**abrupt** scene change) or gradually (**gradual** transitions). Through the use of test metrics, scene changes in video sequences can be detected.

Test metrics are computed by comparing neighboring frames from a video sequence. The nine existing test metrics considered were selected due to the fact that they were not based on object motion. They can be classified into the three main categories, Histogram-based methods, Statistics-based methods, and Template-based methods (as shown in the Table 1).

Histogram-based methods deal with scene lighting distribution. The two metrics studied for this category were the Chi-Square and Absolute Value of Histogram Difference. A numeric representation of the brightness and color in a scene based on the frequency of occurrence of each value is then computed using the formulas in Table 1. Hence, the histogram distribution changes when new RGB values are introduced into the image. The advantage of using Histogram-based methods is the ease of calculation.

Statistics-based methods are based on the first and/or the second order statistics of the image intensity. It consists of the test metrics Liklihood Ratio, F-Test, η_1 , η_2 , and

5

 η_3 . These statistics are calculated with the corresponding formulas in Table 1. Mean (μ) and standard deviation (σ) of the image intensity are used for the analysis.

Template-based methods compare the structure of the adjacent images. The metrics evaluated in this sub-category include Template Matching and Inner Product. Changes in position or movement of objects between the two consecutive frames are easy to realize using this method. Table 1 Summary of Test Metrics

Test Metric	Formula					
Histogram-based method						
$\begin{array}{c} \textbf{Chi-Square} \\ x^2 \ge 0 \end{array}$	$x^{2} = \sum_{i=1}^{M} \frac{\{h_{j}(i) - h_{k}(i)\}^{2}}{\{h_{j}(i) + h_{k}(i)\}}$					
Absolute Value of Histogram Different $0 \le \delta \le 1$	$\delta = \frac{\sum_{i=1}^{M} h_{j}(i) - h_{k}(i) }{\sum_{i=1}^{M} \{h_{j}(i) + h_{k}(i)\}}$					
Statistics-based method						
Likelihood Ratio $\lambda \ge 1$	$\lambda = \frac{\left[\frac{\sigma_j + \sigma_k}{2} + \left(\frac{\mu_j - \mu_k}{2}\right)^2\right]^2}{\sigma_j * \sigma_k}$					
$F-Test$ $F \ge 1$	$F = \frac{\sigma_j^2}{\sigma_k^2}, \sigma_j > \sigma_k$					
$\eta_1 \\ 0 \le \eta_1 \le 1$	$\eta_1 = \frac{\left \mu_j - \mu_k\right * \left \sigma_j^2 - \sigma_k^2\right }{\sigma_j \sigma_k \left(\frac{\mu_j + \mu_k}{2}\right)}$					
$\begin{array}{c} \eta_2 \\ \eta_2 \geq 1 \end{array}$	$\eta_2 = \frac{\mu_j \sigma_j^2}{\mu_k \sigma_k^2}, \mu_j > \mu_k, \sigma_j > \sigma_k$					
η_3 $\eta_3 \ge 1$	$\eta_3 = \left(\frac{\mu_j \sigma_j}{\mu_k \sigma_k}\right), \mu_j > \mu_k, \sigma_j > \sigma_k$					
Template-b:	ased method					
Template Matching $0 \le \Delta \le (255 \times pixels)$	$\Delta = \sum_{x} \left I(x, y, j) - I(x, y, k) \right $					
Inner Product 0≤γ≤1	$\gamma = 1 - \frac{\vec{I}_j \bullet \vec{I}_k}{\left\ \vec{I}_j\right\ \bullet \left\ \vec{I}_k\right\ }$					

7

~

CHAPTER III

ENTROPY-BASED 2D IMAGE DISSIMILARITY MEASURE

Based on the recently proposed Compression-based Dissimilarity Measure (CDM) [1], we propose an Entropy-based Dissimilarity Measure (EDM) as follows

$$EDM(A,B) = \frac{E(A) + E(A - B)}{E(A) + E(B)}$$

Where $E(\bullet)$ is the entropy of a given image, and $\tilde{E}(A - B)$ is the modified entropy of the difference of two images. The difference of two images will increase the dynamic range of the sample; instead of ranges between 0 and 255 for an 8-bit image it will have ranges from -255 to 255. By removing the occurrence of zero values and small difference values (which indicates that two correspondent pixels are the same or their difference is small), the modified entropy of A - B can deal with noise in the image to some degree, and embed the spatial relationship into our dissimilarity metric. When the two images are identical, the modified entropy will return a value zero, and our measure will always have the lower bound value 0.5. However, when the difference image closes to random noises, the $\tilde{E}(A - B)$ may become bigger than E(B), and our measure can have value larger than 1, which is the intuitive upper bound of CDM metric. In general, the smaller the EDM(A, B), the more closely similar the two images A and B are, which is consistent with the behavior of CDM metric.

CHAPTER IV

EXPERIMENTAL DATA AND RESULTS

The proposed metric was tested on a set of still images and videos recoded from CNN TV news channel. We also compared the proposed metric with several existing metrics, i.e. two histogram-based methods, five statistic-based methods, and two template-based methods, that had been used for scene change detection. The formulas for these metrics are shown in Table 1. All the input images were converted from RGB channels to YUV channels first and then the Y-channel was used for testing.

STILL FRAME COMPARISONS

First, the metrics were tested using a conjugate pair of black and white images, as shown in Figure 1. The metrics that based on histogram and statistics methods will consider these two images are the same. The template-based metric and the proposed entropy-based method both will indicate these two images are different to each other. The resultant values from each metric are shown in the second column of Table 2. Values enclosed by a square bracket indicated that the metric considered the two images are the same.



Figure 1 Conjugate pair of black and white images

Second, a set of five images (as shown in Figure. 2) was used for testing dissimilarity measure. Figure 2(a) and 2(b) are two adjacent frames from the same video clip with the person in the scene only moved slightly. Figure 2(c) is a frame from the same clip but with the person moved out the scene. Image shows in Figure 2(d) has some similar background (trees and blue sky). Figure 2(e) is an indoor scene which looks very different as compared with other images. The results of comparing Figure 2(a) with other

images are shown in Table 2. The 3rd column in the Table 2, which we compared Figure 2(a) with itself, is just for insanity check, and all metrics do show the smallest dissimilarity values of their range. All the results are consistent with human subjective judgments, except for one case in the likelihood ratio metric, it considered the Figure 2(e) is more similar to Figure 2(a) instead of Figure 2(d).



Figure 2 Still images

Table 2 Still Images Comparison

	BW	(a) (a)	(a)(b)	(a)(c)	(a)(d)	(a)(e)			
Histogram-based method									
$Chi-Square x^2 \ge 0$	[0]	[0]	1345.75	6700.31	335197	451122			
Histogram Different 0≤δ≤1	[0]	[0]	0.033776	0.0775492	0.652674	0.775648			
Statistics-based method									
Likelihood Ratio λ≥1	[1]	[1]	1.00273	1.00195	8.58188	3.42723			
$F-Test$ $F \ge 1$	[1]	[1]	1.00349	1.00901	1.25798	8.58824			
$\eta_1 \\ 0 \le \eta_1 \le 1$	[0]	[0]	0.000008	0.000017	0.018192	0.099352			
η_2 $\eta_2 \ge 1$	[1]	[1]	1.001235	1.007096	1.162266	8.264914			
η_3 $\eta_3 \ge 1$	[1]	[1]	1.008022	1.005185	1.073837	7.953762			
		Tem	plate-based	method					
Template Matching $0 \le \Delta \le 255(pixels)$	1958400	[0]	228717	255257	357400	4758177			
Inner Product $0 \le \gamma \le 1$	1	[0]	0.002221	0.003377	0.138370	0.931046			
Proposed Entropy (compression) based method									
$Entropy \\ 0.5 \le E \le (1)$	1	[0.5]	0.673985	0.679976	1.01584	1.15243			

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

۰.

VIDEO SCENE CHANGE DETECTION

VIDEO SEQUENCE 1

A short video clip recorded from CNN headline news was selected for testing the proposed dissimilarity measure along with other metrics. The dissimilarity metrics were applied to all the adjacent frames of the video, and the resultant values were normalized to the range between 0 and 1 for the purpose of comparison.

Figure 3, 4, and 5 show sample frames from the recoded CNN headline news. Figure 3 show three consecutive frames (no. 16, 17, and 18) from the video. There is a camera flash in frame number 17, and caused an **abrupt** change in the scene but however the actual scene did not change. This kind of abrupt change is easily detected by all the metrics (as shown in the Table 3). The changes from frame 42 to frame 44 (as shown in Figure 5) could not be easily detected by the histogram-based and statistics-based metrics. However, it had clearly shown up in the plots of the proposed entropy-based metric and template-based metric. Also the **gradual** transitions from frame 21 to 41 (as shown in Figure 4) were clearly shown in plot of the proposed metric without the unwanted dip at frame 33 as shown in the plot of template-based metric.

15



Figure 3 Frames 16 to 18 from video sequence 1



Figure 4 Frames 31 to 34 from video sequence 1



Figure 5 Frames 42 to 44 from video sequence 1

Table 3 Dissimilarity Plot for Each Metric for Video Sequence 1



19



VIDEO SEQUENCE 2

Another video clip recorded from the Weather Channel Headline News was selected for testing the proposed dissimilarity measure along with other metrics. The dissimilarity metrics were applied to all the adjacent frames of the video, and the resultant values were normalized to the range between 0 and 1 for the purpose of comparison. The results are shown in Table 4.

The following figures show sample frames from the recorded Weather Channel Headline News. Figure 6 shows the **gradual** transitions from frames 9 to 20. This kind of change is easily detected by all the metrics (as shown in Table 4). Figure 7 shows the zoom in for frames 50 to 57. This change could not be easily detected by the statisticsbased metrics. Figure 8, 9, and 10 show **abrupt** changes in the scene. These changes could not be easily detected by the statistics-based metrics either. Figure 11 and 12 show the **gradual** transitions from the video. The changes had clearly shown up in the plots of the proposed entropy-based metric, histogram-based metric, and template-based metric.









Figure 7 Frames 50 to 57 from video sequence 2



Figure 8 Frames 60 to 63 from video sequence 2



Figure 9 Frames 156 to 159 from video sequence 2



Figure 10 Frames 205 to 207 from video sequence 2



Figure 11 Frames 288 to 297 from video sequence 2



Figure 12 Frames 350 to 355 from video sequence 2



Table 4 Dissimilarity Plot for Each Metric for Video Sequence 2



31

VIDEO SEQUENCE 3

A short video clip recorded from traffic on a highway was selected for testing the proposed dissimilarity measure along with other metrics. The dissimilarity metrics were applied to all the adjacent frames of the video, and the resultant values were normalized to the range between 0 and 1 for the purpose of comparison.

In this video sequence, the background remained the same through all frames. Scene changes occurred as cars drove past the camera. Figures 13 to 20 shows those frames where significant changes of the dissimilarity measure were shown in the plots. This kind of scene change is easily detected by most of the metrics except the Likelihood ratio and η_1 . The results are shown in the Table 5.



Figure 13 Frames 23 to 28 from video sequence 3



Figure 14 Frames 32 to 37 from video sequence 3



Figure 15 Frames 59 to 66 from video sequence 3





Figure 16 Frames 97 to 110 from video sequence 3



Figure 17 Frames 148 to 155 from video sequence 3

Figure 18 Frames 181 to 188 from video sequence 3

Figure 19 Frames 203 to 214 from video sequence 3

Figure 20 Frames 270 to 275 from video sequence 3

 Table 5 Dissimilarity Plot for Each Metric for Video Sequence 3

44

CHAPTER IV

EXPERIMENTAL DATA AND RESULTS

The proposed metric was tested on a set of still images and videos recorded from CNN TV News Channel. We also compared the proposed metric with several existing metrics, i.e. two histogram-based methods, five statistic-based methods, and two template-based methods, that had been used for scene change detection. The formulas for these metrics are shown in Table 1. All the input images were converted from RGB channels to YUV channels first and then the Y-channel was used for testing.

REFERENCES

- [1] E. Keogh, S. Lonardi, and C. Ratanamahatana, "Towards Parameter-Free Data Mining," in Proc. 10th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, Seattle, Aug. 22-25, WA, 2004.
- [2] M. Li, X. Chen, X. Li, B. Ma, and M. B. Vitánya, "The Similarity Metric," in Proc.
 14th ACM-SIAM Symposium on Discrete Algorithm, Baltimore, MD, 2003, pp. 863 –
 872.
- [3] H. J. Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic Partitioning of Full Motion Video," *Multimedia Systems*, vol. 1, June 1993, pp. 10 – 28.
- [4] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances," in Proc. of the IFIP TC2/WG 2.6 Second Working Conference on Visual Database Systems II, pp. 113 – 127, 1992.
- [5] R. Jain, R. Kasturi, and B. G. Schunck, Machine Vision, McGraw Hill, 1995.
- [6] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes in C, 2dn ed., Cambridge University Press, 1992.
- [7] R. M. Ford, C. Robson, D. Temple, and M. Gerlach, "Metrics for Scene Change Detection in Digital Video Sequences," in *Proc. of the 1997 IEEE Intl. Conf. on Multimedia Computing and Systems (ICMCS '97)*, Ottawa, Ont., June 3 – 6, 1997, pp. 610-611.

- [8] C. Taskiran and E. J. Delp, "Video Scene Change Detection Using The Generalized Sequence Trace," in Proc. of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'98), vol. 5, pp. 2961 – 2964, May 1998.
- [9] ISO/IEC 15444-3, "Information Technology JPEG2000 Image Coding System, Part 3: Motion JPEG2000," 2002.
- [10] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL, 1949.
- [11] J. H. Meng, Y. J. Juan, S. F. Chang, Scene Change Detection in a MPEG Compressed Video Sequence, IS&T/SPIE Symposium Proceedings Vol. 2419, February 1995, San Jose, California
- [12] B. Shahraray, "Scene change detection and content-based sampling of video sequences," in Proc. SPIE'95, Digital Video Compression: Algorithm and Technologies, vol. 2419, San Jose, CA, 1995.

VITA

Meng-Hung Wu

6300 N. 11th st. #43 McAllen, Texas 78504

Education:

B.S. in Atmospheric Sciences National Central University (09/1996—06/2000) Taiwan

Work Experience:

- 2004 2005 Teaching Assistant, University of Texas Pan American, Edinburg, TX
- 2001 2001 Center Administrator/ Mathematics Teacher, Mr. May Education Center, Taiwan
- 2000 2001 Lieutenant in the army, Taiwan
- 1996 2000 Private Teacher, Taiwan

Publications:

Ping-Sing Tsai and Meng-Hung Wu, "Entropy-Based 2D Image Dissimilarity Measure," in proc. of the IEEE international workshop on multimedia signal processing (IEEE MMSP 2005), Shanghai, China, October 30 – November, 2005.