A NEW IMAGE SIMILARITY MEASURE BASED ON ORDINAL CORRELATION

Bogdan Cramariuc[1], *Ilya Shmulevich*[1], *Moncef Gabbouj*[1], *Asko Makela*[2]

[1]Signal Processing Laboratory, Tampere University of Technology
P.O. Box 553, Tampere, Finland
{crama,ilya,moncef}@cs.tut.fi
[2]The Finnish Museum of Photography, Helsinki, Finland
asko.makela@fmp.fi

ABSTRACT

We propose an ordinal-based measure of image similarity. This measure is based on a recently developed general framework for image correspondence and incorporates region-based spatial information. The measure is capable of taking into account differences between images at various scales. Several examples are presented and the measure is evaluated on a set of test images.

1. INTRODUCTION

In many contexts, such as stereo matching [1], image retrieval [2], and motion estimation, a measure of association or similarity between two images is needed. It is desirable for such a measure to be robust in the presence of outliers while being invariant under reasonable image transformations. Outliers are generally caused by various kinds of noise, such as impulsive and bit-error noise. Additionally, different lighting conditions and camera calibrations can distort image intensity values. Thus, it is natural to expect that making an image brighter or even applying a monotonically increasing function to its intensity values should not alter its similarity with another image. Because of such considerations, traditional image matching measures such as correlation or squared Euclidean distance, which are essentially based on pixel intensity values, fail to satisfy our requirements. Indeed, it is well known that the correlation coefficient as well as the squared Euclidean distance are sensitive to outliers and nonlinear monotonically increasing transformations [3].

In order to overcome these problems, so-called ordinal measures of association [4], [5] have been used. The ordinal measures operate on the ranks of pixels rather than directly on the pixel values. Thus, only relative ordering between data values is of consequence in determining the distance or correlation between two images. Two well-known ordinal measures, often used in psychological measurements, are Kendall's τ and Spearman's ρ [5].

Another ordinal correlation measure, κ, was introduced in [6] and [3] for image correspondence. Informally speaking, this measure is based on the ranking of one image with respect to the ranks of the other. This measure is robust in the presence of outliers, is invariant under monotonically increasing transformations, and insensitive to rank distortions [3].

A general framework for performing ordinal-based image correspondence was proposed in [7]. It was shown that this framework contains Kendall's τ and Spearman's ρ as special cases. However, it also allows one to design other correspondence measures that can potentially incorporate region-based spatial information. In the present paper, we consider one such region-based correspondence measure and evaluate its performance on a set of test images. The general structure of the image correspondence framework proposed in [7] is briefly reviewed in the following section.

2. THE IMAGE CORRESPONDENCE FRAMEWORK

Figure 7 gives a general overview of the region-based approach to ordinal image correspondence. Suppose we have two images, X and Y, of equal size. In a practical setting, images of different sizes can be resized and/or reshaped by an appropriate application dependent method. Let $\{X_1, \cdots, X_n\}$ and $\{Y_1, \cdots, Y_n\}$ be the pixels belonging to image X and Y, respectively. We select a number of areas $\{R_1, R_2, \cdots, R_m\}$ and extract the pixels from both images that belong to these areas. Consequently, R_j^X and R_j^Y contain the pixels from image X and Y, respectively, which belong to area R_j.

The goal is to compare the two images using a region-based approach. To this end, we will be comparing R_j^X and R_j^Y, for each $j = 1, \cdots, m$. Because of the considerations mentioned in the Introduction, our approach is an ordinal one and hence, only the ranks of the pixels are to be utilized. For every pixel X_k, we construct a so-called *slice* which is defined as $S_k^X = \left\{ S_{k,l}^X : l = 1, \cdots, n \right\}$, where

$$S_{k,l}^X = \begin{cases} 1, & \text{if } X_k < X_l \\ 0, & \text{otherwise} \end{cases} .$$

As can be seen, slice S_k^X corresponds to pixel X_k and is a binary image of size equal to image X. Slices are built in a similar manner for image Y as well.

With the goal of comparing regions R_j^X and R_j^Y, we first combine the slices from image X, corresponding to all the pixels belonging to region R_j^X. The slices are combined using the operation $OP_1(\cdot)$ into a so-called *metaslice* M_j^X. More formally, $M_j^X = OP_1\left(\left\{S_k^X : X_k \in R_j^X\right\}\right), j = 1, \cdots, m$. Similarly, we combine the slices from image Y to form M_j^Y. It should be noted that the metaslices are equal in size to the original images and could be multi-valued, depending on the operation $OP_1(\cdot)$. Each metaslice represents the relation between the region it corresponds to and the entire image itself.

The next step is a comparison between all pairs of metaslices M_j^X and M_j^Y by using operation $OP_2(\cdot)$, resulting in the *metadifference* D_j. That is, $D_j = OP_2\left(M_j^X, M_j^Y\right), j = 1, \cdots, m$. We thus construct a set of metadifferences $D = \{D_1, D_2, \cdots, D_m\}$. The final step is to extract a scalar measure of correspondence from set D, using operation $OP_3(\cdot)$. In other words, $\lambda = OP_3(D)$. In [7], it was shown that this structure can be used to model the well-known Kendall's τ and Spearman's ρ measures.

3. THE PROPOSED IMAGE SIMILARITY MEASURE

As discussed in the previous section, both images, X and Y, are partitioned into a number of regions. This partition can be obtained by segmenting one of the two images or simply by splitting them into blocks of equal size. Thus, each block in one image is compared to the corresponding block in the other image in an ordinal fashion. Operation OP_1 is chosen to be the component-wise summation operation; that is, metaslice M_j is the summation of all slices corresponding to the pixels in block j or in other words, $M_j^X = \sum_{k : X_k \in R_j^X}\left(S_k^X\right)$. Thus, metaslice j represents the "ranking" of block j with respect to the entire image. Consider two different images shown in Figure 1, taken from [8]. One of the blocks (regions) is indicated on both images by a white square. The size of the blocks is 4×4 pixels. The corresponding metaslices are shown in Figure 2. Note that for the left image, the chosen block contains dark pixels. Most of the image pixels are brighter than all of the pixels in the block, which is why most of the corresponding metaslice is white. For the right image, the selected block contains a larger variety of pixel intensities. This is why there is a greater variation of gray levels in the corresponding metaslice.

Figure 1: Two different images (block size 4×4).

Figure 2: The metaslices corresponding to the blocks shown in Figure 1.

Next, operation OP_2 is chosen to be the squared Euclidean distance between corresponding metaslices. That is, $D_j = \left\|M_j^X - M_j^Y\right\|_2^2$. Figure 3 shows the image produced by operation OP_2 applied to all the metaslices. Note how

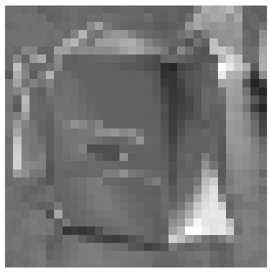


Figure 3: Metadifference image (block size 4×4).

the bright regions in the metadifference image correspond to the regions in which the original images differ. Finally, operation OP_3 sums together all metadifferences to produce $\lambda = \sum_j D_j$. This measure is capable of taking into account differences between images at a scale related to the chosen block size.

To illustrate the effect of the block size, consider the same images with a block size of 32×32, as shown in Figure 4. The corresponding metaslices are shown in Figure 5. Finally, the metadifference image, which is of size 4×4, is shown in Figure 6. Again, it can be seen from the metadifference image that the brighter regions correspond to regions in which the original images differ most.

Figure 4: Two different images (block size 32×32).

Figure 5: The metaslices corresponding to the blocks shown in Figure 4.

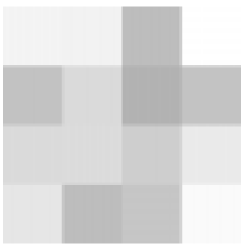


Figure 6: Metadifference image (block size 32×32).

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we illustrate the use of the proposed similarity measure for image matching. We present the results for a set of 13 images taken from The Finnish Museum of Photography database. The images were all resized to 128×128 pixels and the block size was chosen to be 8×8. The value λ was calculated for each pair of images. The results can be seen in Figure 8. There are 6 portraits and 7 landscapes in the test set. As can be seen, portraits are more similar to portraits than to landscapes, and vice versa. The diagonal elements are zero. The gray level in the table indicates the similarity value – darker colors represent more similar images. While these values are not normalized, as is the case for correlation measures, this is not a shortcoming for image matching applications, since only the relative values are relevant.

Future work should focus on choosing meaningful regions by using image segmentation algorithms. Moreover, the effects of dithering and noise on the proposed measure should be studied. Additionally, the image correspondence framework should be generalized to operate on color images. Finally, we believe that there is no one tool that is superior to others in all possible situations. However, a successful framework should be sufficiently flexible in order to accommodate a variety of image types, constraints, and requirements.

Acknowledgment: The authors are grateful to The Finnish Museum of Photography for providing them with a collection of photographs.

5. REFERENCES

[1] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: theory and experiment," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 9, pp. 920-932, 1994.

[2] E. Vicario (Ed.), *Image Description and Retrieval,* Plenum, 1998.

[3] D. N. Bhat and S. K. Nayar, "Ordinal measures for image correspondence," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 4, pp. 415-423, 1998.

[4] R. A. Gideon and R. A. Hollister, "A rank correlation coefficient," *J. Am. Statistical Assoc.*, Vol. 82, No. 398, pp. 656-666, 1987.

[5] M. Kendall and J. D. Gibbons, *Rank Correlation Methods*, fifth ed. New York: Edward Arnold, 1990.

[6] D. N. Bhat and S. K. Nayar, "Ordinal measures for visual correspondence," Columbia Univ., Computer Science, tech. rep. CUCS-009-96, Feb. 1996.

[7] I. Shmulevich, B. Cramariuc, M. Gabbouj, "A framework for ordinal-based image correspondence," *X European Signal Processing Conference (EUSIPCO-2000)*, 5-8 September 2000, Tampere, Finland.

[8] http://www-syntim.inria.fr/syntim/analyse/images-eng.html

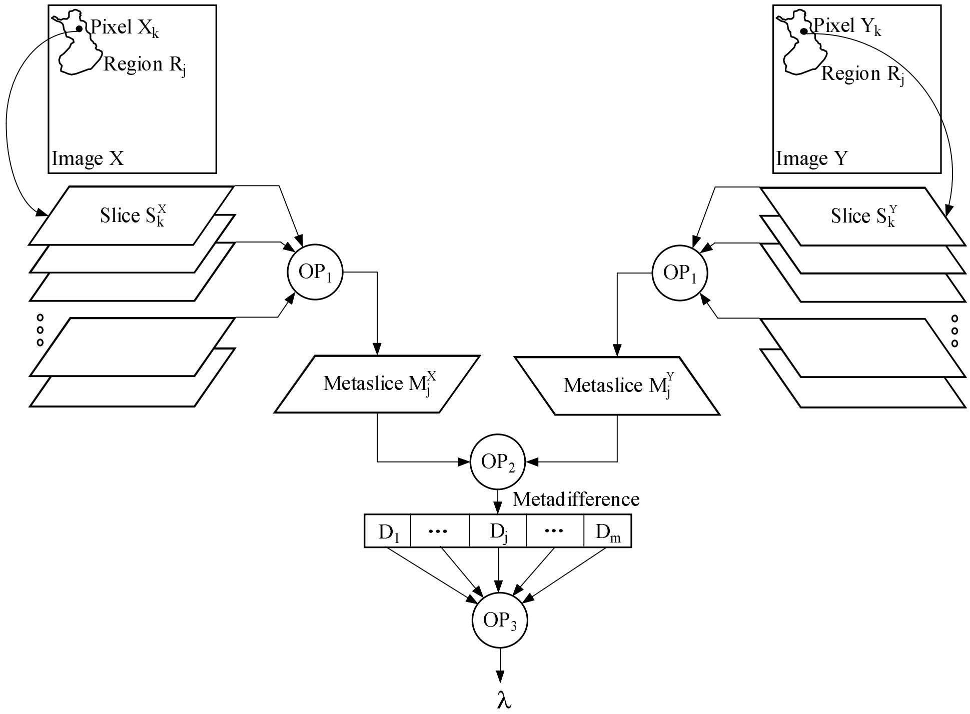
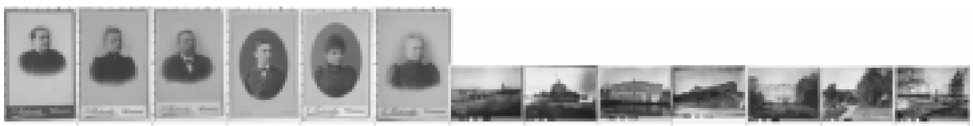


Figure 7: A general framework for ordinal-based image correspondence.



	0.42	0.44	0.65	0.62	0.55	0.74	0.7	0.67	0.74	0.78	0.68	0.71
0.42		0.21	0.39	0.32	0.28	0.83	0.82	0.71	0.66	0.89	0.58	0.75
0.44	0.21		0.37	0.35	0.29	0.74	0.74	0.64	0.61	0.8	0.51	0.7
0.65	0.39	0.37		0.21	0.33	0.73	0.75	0.63	0.57	0.73	0.51	0.64
0.62	0.32	0.35	0.21		0.29	0.8	0.76	0.68	0.59	0.81	0.55	0.69
0.55	0.28	0.29	0.33	0.29		0.63	0.66	0.54	0.46	0.69	0.47	0.64
0.74	0.83	0.74	0.73	0.8	0.63		0.25	0.35	0.38	0.32	0.51	0.34
0.7	0.82	0.74	0.75	0.76	0.66	0.25		0.4	0.39	0.38	0.55	0.4
0.67	0.71	0.64	0.63	0.68	0.54	0.35	0.4		0.42	0.38	0.55	0.48
0.74	0.66	0.61	0.57	0.59	0.46	0.38	0.39	0.42		0.47	0.42	0.45
0.78	0.89	0.8	0.73	0.81	0.69	0.32	0.38	0.38	0.47		0.44	0.35
0.68	0.58	0.51	0.51	0.55	0.47	0.51	0.55	0.55	0.42	0.44		0.5
0.71	0.75	0.7	0.64	0.69	0.64	0.34	0.4	0.48	0.45	0.35	0.5	

Figure 8: Correspondence values λ. All values are divided by 10^9.