ON HISTOGRAMS AND SPATIOGRAMS - INTRODUCTION OF THE MAPOGRAM

Mikael Nilsson, Josef Ström Bartůňek, Jörgen Nordberg, and Ingvar Claesson

Blekinge Institute of Technology
School of Engineering
Box 520, SE-372 25 Ronneby, Sweden
E-mail: mkn@bth.se, jsb@bth.se, jno@bth.se, icl@bth.se

ABSTRACT

This paper introduces the concept of a mapogram. A mapogram may be viewed as a special form of spatiogram, which is a histogram containing additional spatial information. Additionally, this paper presents theory relevant to the creation of a proposed mapogram. A similarity measure derived from the Bhattacharyya coefficient is obtained in order to make comparisons between mapograms. Examples using a mapogram are given.

Index Terms— Image matching, Bhattacharyya coefficient

1. INTRODUCTION

Comparing images, or image patches, is the desired goal in several computer vision applications. The task of building representations or models for an image can be approached in several ways: for example, by Active Appearance Models (AAMs), 2D Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs) [1, 2, 3]. Two other well established, and less computationally expensive, methods used to form a representation involve the use of a template or a histogram. While the template retains exact spatial correspondence, the histogram discards all spatial information. Hence, the two techniques lie at the opposing ends of the spatial information spectrum. Some applications may require spatial information to a greater or lesser degree. Hence, a spatiogram was proposed in order to produce a model somewhere in between a template and a histogram [4].

A spatiogram is basically a histogram with additional spatial information. The potential image comparison capabilities of the spatiogram has been indicated in the field of tracking and video retrieval [4, 5]. Further, a modification to the similarity measure for a spatiogram was proposed in [6].

This paper proposes a different method which incorporates spatial information into a histogram. The spatial information will be contained in 2D probabilistic “maps”. Further, it will be possible to control the amount of spatial information desired in the model.

The paper is organized as follows. The next section presents a review on histograms and spatiograms. Section 3 introduces the theory regarding the proposed mapogram. Section 4 utilizes the mapogram in examples. Finally, conclusions and future work are presented.

2. REVIEW OF HISTOGRAMS AND SPATIOGRAMS

A digital image, in general, is a two dimensional mapping $I : \mathbf{x} \rightarrow v$ from $M \times N$ pixels¹ $\mathbf{x} = [i, j]^T$ to values v . In this work the values v may be arbitrary as long as they can be enumerated. Hence, v can take the form of one of $b = 1, 2, \dots, B$ different discrete values. Often the values v are discrete intensity values in the range $[0 - 255]$; that is, $I(\mathbf{x})$ is a gray-scale image. Among other possibilities for v are color quantization indices (H8S4V2), quantized DCT coefficients, Local Binary Patterns (LBP), local SMQT features, discrete outputs from a local classifier, etc [5, 7, 8, 9, 10].

One method of designing a descriptor of an image is to calculate its histogram. The histogram of an image can be found as

$$n_b = \sum_{i=1}^M \sum_{j=1}^N \delta_b(i, j), \quad \forall b = 1, 2, \dots, B \quad (1)$$

where $\delta_b(i, j) = 1$ if the value v at pixel location $[i, j]$ falls in bin b , and $\delta_b(i, j) = 0$ otherwise. Similarities between two different histograms, n_b and n'_b , can be computed using various techniques, for example, histogram intersection or the Bhattacharyya coefficient. In this paper we adopt the Bhattacharyya coefficient

$$\rho_b^h(n_b, n'_b) = \frac{\sqrt{n_b n'_b}}{\sqrt{\left(\sum_{j=1}^B n_j\right) \left(\sum_{j=1}^B n'_j\right)}} \quad (2)$$

as the similarity measure.

¹Note that the coordinate system follows matrix notation throughout the paper. That is $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$ corresponds to the y- and x-axis, respectively.

In order to extend the histogram with additional spatial information, a spatiogram can be created [4]. A spatiogram adds a mean

$$\mu_b = \frac{1}{\sum_{i=1}^M \sum_{j=1}^N \delta_b(i, j)} \sum_{i=1}^M \sum_{j=1}^N [i, j]^T \delta_b(i, j) \quad (3)$$

and a covariance matrix

$$\Sigma_b = \frac{1}{\sum_{i=1}^M \sum_{j=1}^N \delta_b(i, j)} \sum_{i=1}^M \sum_{j=1}^N ([i, j]^T - \mu_b)([i, j]^T - \mu_b)^T \delta_b(i, j) \quad (4)$$

for each bin as the spatial representation. A number of methods have been proposed to modify the Bhattacharyya coefficient in order to incorporate spatial information (see [4, 6]).

3. INTRODUCTION OF THE MAPOGRAM

One drawback of representing the spatial information by means of a single Gaussian is the fact that multi modal spatial patterns can not be well modeled. This was pointed out in the original spatiogram paper in which it was also mentioned that a possible solution to the problem could be to combine the spatiogram with a Gaussian Mixture Model (GMM) [4]. However, the use of a GMM always provokes questions regarding the number of mixtures to use.

In this paper we propose a different method of incorporating spatial information into a histogram. Rather than calculating a parametric model for each bin, we intend to retain the 2D “map” $\delta_b(i, j)$. Hence the name “mapogram”. Note that a normalized map

$$m_b(i, j) = \frac{\delta_b(i, j)}{\sum_{i=1}^M \sum_{j=1}^N \delta_b(i, j)} \quad (5)$$

is a 2D discrete probability function. If $m_b(i, j)$ were to be applied directly as in Eq. (5), the spatial information would imply an exact spatial correspondence. In other words, the only sample that can be drawn from a mapogram created from a single image would be the exact same image. On the other hand, if

$$m_b(i, j) = \frac{1}{MN} \quad (6)$$

then the mapogram would be exactly the same as a standard histogram.

In order to control the amount of spatial information a mapogram should represent, a parameter $0 \leq \alpha \leq 1$ is introduced. A general map, incorporating α , is denoted $\mathbf{m}_{b,\alpha}$. If $\alpha = 0$ then the exact spatial map $\mathbf{m}_{b,0}$ is employed and

found according to Eq. (5). Similarly, a mapogram using a map $\mathbf{m}_{b,1}$ found according to Eq. (6) neglects all spatial information. Further, the Bhattacharyya similarity measure between two maps can be found for every bin as

$$\rho_b^m(\mathbf{m}_{b,\alpha}, \mathbf{m}'_{b,\alpha}) = \sum_{i=1}^M \sum_{j=1}^N \rho_{b,i,j}^s \quad (7)$$

where

$$\rho_{b,i,j}^s(m_{b,\alpha}(i, j), m'_{b,\alpha}(i, j)) = \frac{\sqrt{m_{b,\alpha}(i, j), m'_{b,\alpha}(i, j)}}{\sqrt{\left(\sum_{i=1}^M \sum_{j=1}^N m_{b,\alpha}(i, j)\right) \left(\sum_{i=1}^M \sum_{j=1}^N m'_{b,\alpha}(i, j)\right)}} \quad (8)$$

Note that the denominator in Eq. (8) can be neglected if the maps are normalized. By combining Eq. (2) and Eq. (7) a similarity measure between two mapograms can be found as

$$\rho_\alpha = \sum_{b=1}^B \rho_b^h(n_b, n'_b) \rho_b^m(\mathbf{m}_{b,\alpha}, \mathbf{m}'_{b,\alpha}). \quad (9)$$

The question is, how to form a map with α -values in between zero and one in order to adjust the amount of spatial information that will affect the distance measure. The two extreme maps $\mathbf{m}_{b,0}$ and $\mathbf{m}_{b,1}$ (as well as a given α) form the input for the design of a general map $\mathbf{m}_{b,\alpha}$.

The first step in the design of a general α -map is to modify $\mathbf{m}_{b,0}$ to reduce the strong spatial correspondence. This is achieved by filtering $\mathbf{m}_{b,0}$ with a 2D blurring kernel that is dependent on α . Here we adopt a Gaussian blur, that is, the kernel is found as

$$K_\alpha(i, j) = \frac{1}{2\pi |\Sigma_\alpha|^{1/2}} e^{-\frac{1}{2}([i, j]^T - \mu)^T \Sigma_\alpha^{-1}([i, j]^T - \mu)} \quad (10)$$

where $\mu = [M/2, N/2]^T$ and

$$\Sigma_\alpha = \begin{bmatrix} \sigma_{y,\alpha}^2 & 0 \\ 0 & \sigma_{x,\alpha}^2 \end{bmatrix} \quad (11)$$

where $\sigma_{y,\alpha}^2 = \alpha \frac{M}{4}$ and $\sigma_{x,\alpha}^2 = \alpha \frac{N}{4}$ control the strength of the Gaussian blur. Note that if $\alpha = 0$ then the Gaussian kernel is replaced with a Dirac delta function and no filtering/blurring occurs.

The found filtered result is a map which expand existing points to the vicinity around them. However, the map will never become completely flat as $\mathbf{m}_{b,1}$. In order to achieve this at $\alpha = 1$, a simple α -weighted average is performed and a final map can be found as

$$\mathbf{m}_{b,\alpha} = (1 - \alpha) [\mathbf{m}_{b,0} * \mathbf{K}_\alpha] + \alpha [\mathbf{m}_{b,1}] \quad (12)$$

where $*$ is the 2D convolution operator. Further, the map $\mathbf{m}_{b,\alpha}$ is normalized to sum up to one (or indirectly normalized in the comparison by utilizing the denominator in Eq. (8)). An example of the effects of the operations performed in Eq. (12) can be found in Fig. 1.

Fig. 1. Example of the map effect for different α values. Note that the original map is exactly the same as $\mathbf{m}_{b,0}$ (full spatial correspondence), and that $\mathbf{m}_{b,1}$ is exactly uniform (the mapogram is equal to a histogram).

In order to create a complete mapogram of an image $I(\mathbf{x})$ of size $M \times N$, the number of values B and the spatial importance parameter α needs to be defined. A calculated mapogram constitutes of a histogram according to Eq. (1) and a map for each bin according to Eq. (12). In general, the histogram and every map are normalized to sum to one. A comparison between two mapograms can then be performed according to Eq. (9).

4. EXAMPLES

In this section three different examples are given in order to present a proof of concept for the proposed mapogram. The images used are mainly face images from the Olivetti Research Laboratory (ORL) face database. In all examples the gray-scale indices are used as features.

4.1. Generate Samples

In this first example the first image found in the ORL database is used to create a mapogram. Randomly generated samples are then extracted and can visually indicate the amount of spatial information represented for each mapogram, see Fig. 2. Note how the spatial information of the faces degrades as α increases.

4.2. Compare Images

Here a face image is compared to seven other images, see Fig. 3. The first image is the original image but with pixels in

Fig. 2. Original image and examples of samples generated from mapograms with different α -values. Note how the spatial information degrades as α increases.

a small square set to zero. The second image is the original image rotated 15 degrees. The third image is the original image upside-down. The fourth is a different image but with the same person. Fifth is a face from another person. The last two images are nonface images; one of a flower and one randomly generated image. The images are compared with mapograms of different α -values, see Fig. 4. Note that the versions of the original image are rated higher than the other person and the nonface images. Also note that around $\alpha = 0.5$ a shift occurs from Image 1 to Image 3 (see Fig. 3) as the top scoring image. This since the spatial information is more neglected at higher α -values, and the histogram section begins to take over the distance score in Eq. (9).

Fig. 3. Original image and seven images for comparison.

4.3. Face Identification

This last example compares different α -values for the task of face identification (1-to-many comparison). Only the first sample per person is used for training; the remaining nine from the ORL database are used for evaluation for all 40 persons. As a reference, the spatiogram proposed in [6] is applied on the same task resulting in a Recognition Rate (RR) of 73.1%. Note that a mapogram with $\alpha = 0.2$ results in the highest RR of 76.9% for this example. However, different tasks, objects and variation may of course require different α -values.

Fig. 4. Order of similarity (top most similar - bottom least similar) for different α -values for the mapogram.

α	0	0.1	0.2	0.3	0.4	0.5
RR	63.6	76.7	76.9	73.6	71.7	68.3
α	0.6	0.7	0.8	0.9	1	
RR	64.7	64.7	63.9	63.3	62.8	

Table 1. Recognition Rate (RR) for different α -values used for the mapogram on the ORL face database. For comparison, the proposed spatiogram in [6] results in a RR of 73.1% on the same task.

5. CONCLUSIONS

This paper presented the concept of a mapogram. A mapogram may be viewed as a histogram with additional spatial information represented by probabilistic maps. The relevant theory was presented including the derivation of the Bhattacharyya similarity measure and examples of the use of the mapogram. As indicated by the examples, one of the key advantages of the mapogram is the possibility to adjust a tunable parameter, α , that controls the amount of spatial information contained in the mapogram representation.

6. FUTURE WORK

Memory is an obvious issue for the proposed mapogram. In some applications with large image size and many discrete values, memory issues may be problematic. Scaling down the size of the maps used may provide a possible solution to this problem. For example, an image of size $M \times N$ may use maps of size $M_r \times N_r$, where $M_r \leq M$ and $N_r \leq N$. The scaling operation must be properly designed for this task, as probabilistic maps are employed and not regular images. Hence, standard image resize functions may not be suitable. Further, the resize operation indirectly implies that there can be no strong spatial correspondence, i.e. α -values close to

zero in the original size can not be used. Therefore a virtual threshold α_{min} ($\alpha_{min} \leq \alpha \leq 1$) is imposed that restricts the choices of α in the full size. The exact details for the map resize operation and the limitations of α in a resized map may be subject of future work.

7. REFERENCES

- [1] Iain Matthews, Takahiro Ishikawa, and Simon Baker, "The template update problem," in *Proceedings of the British Machine Vision Conference*, September 2003.
- [2] A.V. Nefian and III. Hayes, M.H., "Maximum likelihood training of the embedded HMM for face detection and recognition," in *IEEE Int. Conf. on Image Processing*, September 2000, vol. 1, pp. 33–36.
- [3] J. Bilmes, "A gentle tutorial on the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models," Tech. Rep., 1997.
- [4] Stanley T. Birchfield and Sriram Rangarajan, "Spatigrams versus histograms for region-based tracking," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2005, pp. 1158–1163.
- [5] Adrian Ulges, Christoph H. Lampert, and Daniel Keyers, "Spatigram-based shot distances for video retrieval," in *Trecvid 2006 Workshop*, 2006.
- [6] Ciaran O Connaire, Noel O'Connor, and Alan F. Smeaton., "An improved spatiogram similarity measure for robust object localisation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2007, pp. 1069–1072.
- [7] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971–987, 2002.
- [8] O. Lahdenoja, M. Laiho, and A. Paasio, "Reducing the feature vector length in local binary pattern based face recognition," in *IEEE International Conference on Image Processing (ICIP)*, September 2005, vol. 2, pp. 914–917.
- [9] M. Nilsson, M. Dahl, and I. Claesson, "The successive mean quantization transform," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2005, vol. 4, pp. 429–432.
- [10] M. Nilsson, J. Nordberg, and I. Claesson, "Face detection using local smqt features and split up snow classifier," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, April 2007.