Computer Sciences Department

Face Cyclographs for Recognition

Guodong Guo
Charles Dyer

Technical Report #1555

March 2006

UNIVERSITY OF
WISCONSIN
M A D I S O N

Face Cyclographs for Recognition *

Guodong Guo and Charles R. Dyer
Computer Sciences Department
University of Wisconsin-Madison
Madison, WI 53706
{gdguo, dyer}@cs.wisc.edu

Abstract

A new representation of faces, called face cyclographs, is introduced for face recognition that incorporates all views of a rotating face into a single image. The main motivation for this representation comes from recent psychophysical studies that show that humans use continuous image sequences in object recognition. Face cyclographs are created by slicing spatiotemporal face volumes that are constructed automatically based on real-time face detection. This representation is a compact, multiperspective, spatiotemporal description. To use face cyclographs for recognition, a dynamic programming based algorithm is developed. The motion trajectory image of the eye slice is used to analyze the approximate single-axis motion and normalize the face cyclographs. Using normalized face cyclographs can speed up the matching process. Experimental results on more than 100 face videos show that this representation efficiently encodes the continuous views of faces and improves face recognition performance over view-based methods.

1 Introduction

Over the last several years there have been numerous advances in capturing multiperspective images, i.e., combining (parts of) images taken from multiple viewpoints into a single representation that simultaneously encodes appearance from many views. Multiperspective images [23, 15] have been shown to be useful for a growing vari-ety of tasks, notably scene visualization (e.g., panoramic mosaics [11] [16]) and stereo reconstruction [14]. Since one fundamental goal of computer vision is object recognition [9], a question may be asked: are multiperspective images of benefit for object recognition?

Under normal conditions, 3D objects are always seen from multiple viewpoints, either from a continuously moving observer who walks around an object or by turning the object so as to see it from multiple sides. This suggests that a multiperspective representation of objects might be useful.

Recently, psychophysical results have shown that the human brain represents objects as a series of connected views [18] [21] [2]. In psychophysical experiments by Stone [18], participants learned sequences which showed 3D shapes rotating in one particular direction. If participants had to recognize the same object rotating in the opposite direction, it took them significantly longer to recognize and the recognition rate decreased. This result cannot be reconciled with traditional view-based representations [19] whose recognition performance does not depend on the order in which images are presented. Instead, it is argued in [18] that the temporal characteristics of the learned sequences such as the order of images are closely intertwined with object representation. These results and others from physiological studies [10] support the hypothesis that humans represent objects as a series of connected views [2].

The findings from human recognition may have practical guidance for developing better computational object recognition systems. Bülthoff et al. [2] presented a method for face recognition based on psychophysical results [18] [21] in which they showed experimentally that

the representation of connected views gives much better recognition performance than traditional view-based methods. The main idea of their approach is to process an input sequence frame-by-frame by tracking local image patches to achieve segmentation of the sequence into a series of time-connected "key frames" or views. However, a drawback of the "key frames" representation is that it still needs several view images to characterize the whole sequence.

Can we integrate all continuous views of an object into a *single* image representation? In this paper we propose a new method to incorporate all views of an object, which is the cyclograph of the object [5] [15], a type of multi-perspective image. A cyclograph is generated when the object rotates in front of a static camera or the camera rotates around the object. See Figure 5 for some examples of face cyclographs. It has a long history in photography, with the first patent related to making cyclographs in 1911 [5]. Cyclographs have been used previously for image-based rendering [13] and stereo reconstruction [14]. But, to our knowledge, there is no previous work using cyclographs for object recognition.

The major contributions of this paper are:

- Propose a new method to incorporate all continuous views of an object for recognition using the cyclograph image of the object.

- Develop a method to recognize faces using cyclographs based on a dynamic programming technique which can align and match cyclographs simultaneously.

- Present a method to normalize cyclographs based on motion trajectory image analysis and image warping. Using the normalized cyclographs a faster method for recognition is developed.

The paper is organized as follows. Section 2 presents the analysis of the *spatiotemporal volume* of continuous views of objects, and the generation of face cyclographs. Section 3 describes properties of face cyclographs especially for face recognition. Section 4 presents two methods for face recognition using face cyclographs. Experimental results are given in Section 5. Some issues are discussed in Section 6.

2 Viewing Rotating Objects

Our goal is to develop a computational method that encodes all continuous views of faces for face recognition. In psychophysical experiments, the connected views of an object are captured from object rotation in one particular direction [18] [2]. Following the psychophysical experiments, we consider the class of single-axis rotations and associated appearances as the basis for capturing the continuous views of faces. The most natural rotations in depth for faces are when an erect person rotates his or her head, resulting in an approximately single-axis rotation about a vertical axis. Many other objects have single-axis rotations as the most "natural" way of looking at them. When we see a novel object we usually do not see random views of the object but in most cases we walk around it or turn the object in our hand [2].

2.1 Spatiotemporal Volume

Suppose that a 3D object rotates about an axis in front of a camera, as shown in Figure 1, where different circles represent different depths of the object, and a sequence of images are captured. Stacking together the sequence of images, a 3-dimensional volume, x-y-t, can be built, which is called a *spatiotemporal volume*. All continuous views are contained within this 3D volume data.
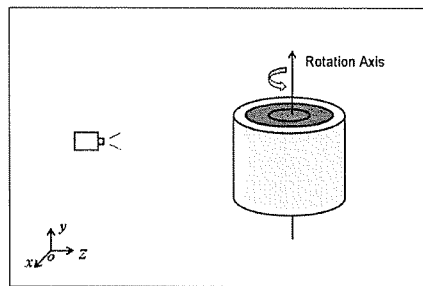

Figure 1: A camera captures a sequence of images when an object rotates about an axis. Circles with different radii denote different depths of the object.

In psychophysical studies, this 3D volume data is called a *spatiotemporal signature* and there is evidence showing that such signatures are used by humans in object recog-

2

nition [17], but no computational representation was presented. We analyze the spatiotemporal volume and generate a computational representation of rotating objects.

2.2 3D Volume Analysis

The *spatiotemporal volume*, x-y-t, is a stack of x-y images accumulated over time t. Each x-y image contains only appearance but no motion information. On the contrary, the x-t or y-t images contain both spatial and temporal information. They are called *spatiotemporal images*. The x-t and y-t images can be obtained by slicing the x-y-t volume, as shown in Figure 2, with blue and red lines, respectively.
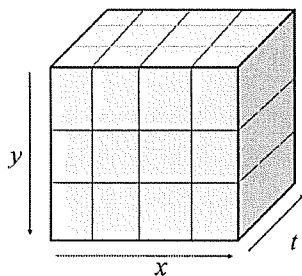

Figure 2: A 3-dimensional volume is sliced to get different image content. The x-t and y-t slices are *spatiotemporal images*.

Given a 3D volume, x-y-t, all the x-t (or y-t) slices preserve all the original information without any loss. This is not difficult to see. The x-y slices are captured by the camera, while the x-t or y-t slices are cut from the volume independently. The union of all x-t (or y-t) slices is exactly the original volume. On the other hand, different slices, i.e., x-y, x-t, or y-t, encode different information from the 3D volume.

Although both x-t and y-t slices are *spatiotemporal images*, they contain different information. When the object rotates about an axis that is parallel to the image's y axis, each x-t slice contains information on object points along a horizontal line on the object surface, defining the motion trajectories of these points. One example is shown in Figure 8(a). On the contrary, each y-t slice contains column-wise appearance of the object surface because of

the object rotation about an axis that is parallel to the image's y axis. Thus y-t slices encode the appearance of the object as it rotates 360°. Partial examples are shown in Figure 7.

When a convex (or nearly convex) object rotates about an axis 360°, the *spatiotemporal volume* is constructed by stacking the whole sequence of images captured by a static camera. The slice that intersects the rotation axis usually contains the most visible appearance of the object in comparison with other parallel slices. Furthermore, this slice has also least distortion.

As shown in Figure 3 with a top-down view, when an object rotates 360°, each point on the object surface intersects the slice, S_4, once and only once. All other slices will miss seeing some parts of the object. In this sense S_4 contains the most appearance of the object. This can also be observed from the y-t slices in the face volume shown in Figure 7 in which the middle image corresponding to S_4. Further, slice S_4 usually minimizes foreshortening distortion because it samples every visible surface point at a near-normal angle. While other parallel slices approach the object surface at variable angles.
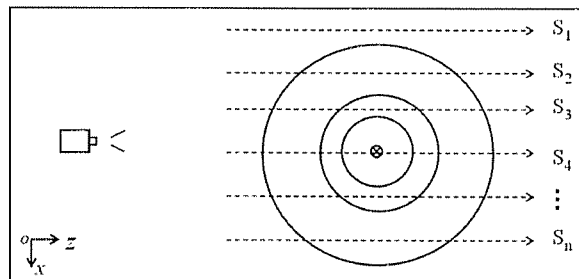

Figure 3: Top-down view of a 3D object rotating about an axis. The circles with different radii denote different depths on the object surface.

2.3 Spatiotemporal Face Volume

To represent rotating faces for recognition we need to extract a spatiotemporal sub-volume containing the face region, which we call the *spatiotemporal face volume*. A face detector [20] can be used to automatically detect faces in sequences of face images. Figure 4 shows the

3

face detection results in the first frame of a face video. The face positions reported by the face detector can then be used to determine a 3D face volume. False alarms from the face detector are removed by using facial skin color information. The eyes, detected with a similar technique as that in the face detector [20], are used for locating the motion trajectory image of the eye-level slice, which will be presented in Section 4.3.

Figure 4: Face and eye detection in a frontal face image.

2.4 Face Cyclographs

Given a *spatiotemporal face volume* with each coordinate normalized between 0 and 1, we can analyze the 3D face volume via slicing. Based on Section 2.2, one may slice the volume in any way without information loss. However, the y-t slices encode all of the visible appearance of the object for single-axis rotation about a vertical axis. Furthermore, the unique slice that intersects the rotation axis contains the most visible appearance of the object with minimum distortion among all y-t slices. As a result, we will use this slice for the rotating face representation.

In our face volume, the slice that intersects the rotation axis is approximately the one with $x = 0.5$. This middle slice extracts the middle column of pixels from each frame and concatenates them to create a face-like image, called the "cyclograph of a face," or simply "face cyclograph." One face cyclograph is created for each face video. The size of a face cyclograph image is determined by the video length and the size of the segmented faces, i.e., the width of the face cyclograph is the number of frames in the video, and the height is the height of the segmented faces.

A face cyclograph can also be viewed as being captured by a strip camera [13]. As shown in Figure 6(b), the face cyclograph captures the face completely from left to right profiles, and all parts of the face surface are captured equally well. On the contrary, when a pin-hole camera is used as shown in Figure 6(a), the face surface is captured poorly when the camera's viewing rays approach grazing angle with the face surface, and the face surface, causing parts of the face surface to be captured unequally.
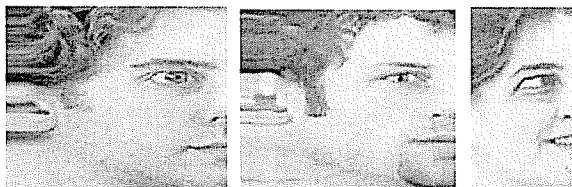

Figure 5: Some examples of face cyclographs. Each head rotates approximately 90°.
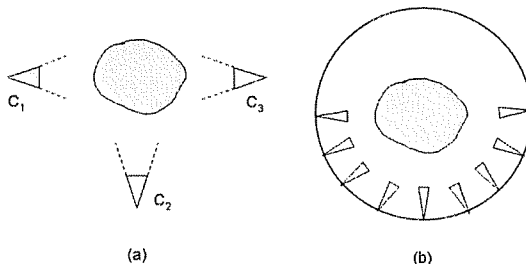

(a)   (b)

Figure 6: A face (nearly-convex object) is captured. (a) The frontal (from C_2) and side views (from C_1 and C_3) are captured separately. (b) The face cyclograph captures all parts of the face surface equally well.

Because in our face videos (see Section 5.1 for details) the initial face is always approximately frontal and the last face is approximately a profile view, the created face cyclographs look like a "half face," as shown in Figure 5. To create a "whole face cyclograph," the head needs to rotate approximately 180°. For recognition purpose, there is no need to capture 360° head rotation since the back of the head has no useful information.

4

Figure 7: The *y-t* slices of the face volume at every twenty-pixel interval in the *x* coordinate.

3 Properties of Face Cyclographs

Some properties of the face cyclograph representation are now described, especially concerning the face recognition problem.

3.1 Compact

The face cyclograph representation is compact. From Section 2, the *y-t* slices contain all appearance information in a *spatiotemporal face volume*. But only one slice intersects the rotation axis (see Figure 3). The face cyclograph is constructed from this slice. The other slices that do not intersect the rotation axis are not used. Consequently, this representation largely reduces the redundancy in the *spatiotemporal face volume*. In comparison with Bülthoff's key frames approach [2], the face cyclograph uses local strips from moving faces without overlap, instead of using partially overlapped key frames and overlapped local patches from each key frame. Therefore the face cyclograph is a concise representation.

3.2 Multiperspective

A face cyclograph is a multiperspective image of a face. The advantage of using a multiperspective face image is that the faces observed from all viewpoints can be integrated together into a single image representation. The multiperspective face image encodes facial appearance all over the face surface and not just from 1 viewpoint. The face cyclograph can be viewed as being captured by a strip camera [13]. For nearly cylindrical objects (e.g., faces), each strip captures frontoparallel views of the surface along that strip. On the contrary, the "key frames" approach [2] uses a series of single perspective images.

3.3 Keeps Temporal Order

If a head rotates continuously in one direction, the face cyclograph successively extracts strips from the spatiotemporal face volume without changing the temporal order in the original face sequence. Temporal order is important for moving face recognition in psychophysical studies [17] [18] [2].

Computationally, the temporal order is also important for designing a matching algorithm for face recognition. In Section 4 the recognition algorithm, which is based on dynamic programming, depends on this property.

4 Recognition using Face Cyclographs

For face recognition, one face cyclograph is computed for each face video sequence containing one rotating face. Given a testing face sequence, the face cyclograph is computed first and then matched to all face cyclographs in the database. Two algorithms have been developed for matching face cyclographs. The first uses dynamic programming (DP) [12] for alignment and matching of face cyclographs. The monotonicity condition has to be satisfied to use DP and face cyclographs satisfy this by keeping the temporal order of the original face sequences. The second method analyzes the face motion trajectory image and then normalizes face cyclographs to the same size before matching.

4.1 Matching Two Strips

The local match measure of two strips is described in this subsection. Each strip is a vertical column in a face cyclograph image. Matching two strips in two face cyclographs is a 1D image matching problem. We define the similarity

5

between two strips, $strip_i$ and $strip_j$, using the α-norm:

$$S_{i,j}^{\alpha} = \| \rho_1(strip_i, \Theta) - \rho_2(strip_j, \Theta) \|_{\alpha} \qquad (1)$$

where ρ_1 and ρ_2 are transforms for strips with respect to a parameter set Θ. Θ characterizes the methods used for feature extraction. Currently, we simply use the pixel color information as the similarity measure; one could alternatively use a 1D wavelet transform to extract features and then match strips.

4.2 Matching Face Cyclographs using Dynamic Programming

Given a match measure between two strips, the next step is to match two face cyclographs. The number of strips within each cyclograph will vary in general because it is determined by the number of frames in the input video sequence, which itself is influenced by the speed and uniformity of the head rotation. The algorithm has to take these variabilities into consideration in matching face cyclographs.

We develop a method for matching face cyclographs based on the dynamic programming technique [12], which can effectively align variable-width face cyclographs and match them simultaneously.

The DP technique can be used for matching face cyclographs because they keep the temporal order in head motion. The sub-problem of matching two strips was presented in Section 4.1.

The cost function of DP optimization is

$$C_{i,j} = \min\{C_{i-1,j-1}, C_{i-1,j}, C_{i,j-1}\} + S_{i,j}^{\alpha} \qquad (2)$$

where $C_{i,j}$ is the minimum cost of matching strip pairs i and j. The accumulated costs are filled in a 2D table and an optimal path is traced back in the cost table. The final cost corresponds to the optimal path to match two face cyclographs. The smaller this cost, the more similar are two face cyclograph images.

The computational complexity of dynamic programming is $O(MN)$ to match two face cyclographs of widths M and N.

4.3 Normalized Face Cyclographs

Face cyclographs can also be normalized to the same size before matching. Using normalized face cyclographs

can make the recognition process much faster, and allow feature extraction on 2D images rather than 1D strips. To normalize face cyclographs, we developed a method based on motion trajectory image analysis.

Motion-trajectory images are slices perpendicular to the rotation axis in the spatiotemporal volume. They are similar to the epipolar plane images (EPI) [1]. The EPI was used for scene structure estimation with a camera moving along a straight line. Here we use the motion trajectory images for face motion analysis. For a face rotating about a vertical axis, the horizontal slices contain face motion trajectory information. Experimentally we found that the slice of the eyes gives richer information than other slices for motion analysis. One example of the eye slice is shown in Figure 8(a).

Given the eye slice motion-trajectory image, we can detect and remove non-motion image frames from the original sequence of face images, and then align the remaining frames. The whole algorithm consists of the following 5 steps:

(1) Edge detection. Edges in the motion trajectory image are detected using the Canny edge detector [3].

(2) Average edge direction. The average of edge directions over each row in the edge image is estimated using

$$\overline{Dir_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} \| \cot \theta_{ij} \|$$

where n_i is the number of edges in row i of the motion trajectory image, θ_{ij} is the edge direction angle of the jth edge in row i, and $\overline{Dir_i}$ is the average of edge direction in row i. This average improves the robustness for edge direction estimation.

(3) Median filtering.

(4) Non-motion detection. Each row in the motion trajectory image corresponds to one frame in the original video sequence. $\overline{Dir_i}$ characterizes the amount of motion in frame i. If the edge in row i is almost vertical then there is no motion in frame i, and the value of $\overline{Dir_i}$ will be very small. So the criterion for non-motion detection is that if $\overline{Dir_i}$ is smaller than a threshold (experimentally chosen to be 0.4), frame i contains no motion. The detected frames with no motion are removed.

(5) Image warping. The remaining frames in the image sequence contain some head rotation between consecutive frames. The corresponding strips sliced from
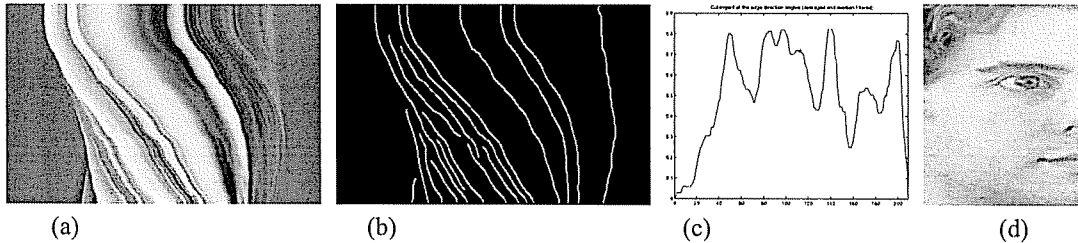
6

Figure 8: (a) Motion trajectory image sliced along the right eye center. (b) Detected edges. (c) Cotangent of the edge direction angles averaged and median filtered. (d) The new face cyclograph after non-motion part removal.

those frames are concatenated to construct the face cyclograph. In this way, all face cyclographs contain only moving parts. Finally, the face cyclograph is normalized to a fixed size by image warping [22].

5 Experiments

5.1 A Dynamic Face Database

A face video database with horizontal head rotation was captured. Each subject was asked to rotate his or her head from an approximately frontal view to an approximately profile view (i.e., approximately a 90° head rotation). A single, stationary, uncalibrated camera was used to capture videos of the subjects. 28 individuals, each with 3 to 6 videos, were captured for a total of 102 videos in the database. The number of frames per video varies, ranging from 98 to 290, resulting in a total of 21,018 image frames. Each image is size 720×480. An image in one of our face videos is shown in Figure 4.

Each video in our face video database was matched against all other face videos, providing an exhaustive comparison of every pair of face videos. Precision and recall measures were computed to evaluate the algorithm's performance. Let TP stand for true positives, FP for false positives, and FN for false negatives. $precision$ is defined as $\frac{TP}{TP+FP}$, and $recall$ is defined as $\frac{TP}{TP+FN}$. Precision measures how accurate the algorithm is in predicting the positives, and recall measures how many of the total positives the algorithm can identify. Both precision and recall were computed with respect to the top n matches, characterizing how many faces have to be examined to get a desired level of performance.

5.2 Face Recognition Results

Face cyclographs were created for all 102 face videos in our database. No faces were missed by this completely automatic process. The similarity measure between two face cyclographs was the 1-norm, i.e., $\alpha = 1$ in Eq. (1). Given a query face cyclograph, the costs of matching it with all remaining 101 face cyclographs were computed and sorted in ascending order. Then the precision and recall were computed with respect to the top n matches, with $n = 1, 2, \cdots, 101$. Finally, the precision and recall were averaged over all 102 queries and are shown in Figure 9. The average precision was very high for the top matches. For example, it was 99.02% for the top 1 match, 95.10% for the top 2, and so on for the DP method. The recall curve approaches 1 very quickly.

Using the normalized face cyclograph method, the performance was slightly lower than using DP. The reason may be that linear warping introduces artifacts. A nonlinear warping method is under consideration.

The face cyclographs algorithms were also compared with a simple view-based face recognition method, similar to those in [2]. The test faces in [2] were the faces between key frames. Since there is no manual labelling in our approach, we randomly chose one frame in each video and did face detection [7] together with false detection removal. We used the same 1-norm image similarity measure similar to Eq. (1) but defined in 2D. As seen in Figure 9, the performance of the face cyclographs methods is much higher than view-based face recognition in both precision and recall.

We did not compare our face cyclographs approach with the "key frames" approach [2] because it requires labelled faces, which are not easily obtained for the amount

7

of video in our database. Our approach uses dense sequences without any manual work.

6 Discussion

In this paper we used face cyclographs for face recognition, which integrated the continuous views of a rotating face into a single image. We believe that this multiperspective representation is also useful for other object representation and recognition tasks. The basic idea is to capture object appearance from a continuous range of viewpoints and then generate a single multiperspective image to represent the object, instead of using multiple single-perspective images which is the traditional view-based representation.

Assuming a simplified 3D head model, e.g., a cylinder [4] or ellipsoid [8], a 2D face image taken from a single viewpoint can be unwrapped when it is registered with the head model that contains reference face texture maps. Our face cyclograph representation does not require any assumptions about the object shape, nor registration of different object views. Hence it is not difficult to extend the cyclograph representation for other object recognition tasks. Furthermore, the creation of a face cyclograph is simple and fast so it is useful for real-time recognition. Finally, unwrapped faces [4] [8] are not necessarily multiperspective [15], as face cyclographs are.

The focus of this paper is to propose a face representation that encodes all views of a rotating face with a face cyclograph, and to demonstrate its use for face recognition. Our work is different from recent methods on video-based face recognition where head motions are arbitrary (see [6] and references there).

In order to extend the face cyclograph representation for a face video containing arbitrary head motions, a key pre-processing step is required. That is, to manipulate a face video with arbitrary head motion to create a face video with single axis head rotation starting from a frontal view and ending at a side view. This pre-processing step can be viewed as an image-based rendering problem. Then, a face cyclograph can be generated and used for recognition based on the techniques presented in this paper. We will investigate this issue in the future.

7 Conclusions

Motivated by recent psychophysical studies, this paper presented a new face representation, called face cyclographs, for face recognition. Temporal characteristics are encoded as part of the representation, resulting in better face recognition performance than using traditional view-based representations. This new representation is compact, robust, and simple to compute from a *spatiotemporal face volume*, which itself is automatically constructed from a video sequence. Face recognition is performed using dynamic programming to match face cyclographs or using normalized face cyclographs based on motion trajectory analysis and image warping. We expect the multiperspective representation to improve results of other object recognition tasks as well.

References

[1] R. C. Bolles, H. H. Baker, and D. H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *Int'l J. of Computer Vision*, 1(1):7–55, 1987.

[2] H. H. Bülthoff, C. Wallraven, and A. Graf. View-based dynamic object recognition based on human perception. *Proc. 16th Int. Conf. Pattern Recognition*, 3:768–776, 2002.

[3] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.

[4] M. Cascia, S. Sclaroff, and V. Athitsos. Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3d models. *IEEE Trans. Pattern Analysis and Machine Intell.*, 22(4):322–336, 2000.

[5] A. Davidhazy. Principles of peripheral photography. In *http://www.rit.edu/ andpph/text-peripheral-basics.html*.

[6] A. Hadid and M. Pietikinen. An experimental investigation about the integration of facial dynamics in video-based face recognition. *Electronic Letters on Computer Vision and Image Analysis*, 5(1):1–13, 2005.

[7] S. Li, L. Zhu, Z. Zhang, A. Blake, H. Zhang, and H. Shum. Statistical learning of multi-view face detection. In *Proc. 7th European Conference on Computer Vision*, volume 4, pages 67–81, 2002.

[8] X. Liu and T. Chen. Pose-robust face recognition using geometry assisted probabilistic modeling. In *Proc. Computer Vision and Pattern Recognition*, pages 502–509, 2005.

8

[9] D. Marr. *Vision*. Freeman Publishers, 1982.

[10] Y. Miyashita. Neural correlate of visual associative long-term memory in the primate temporal cortex. *Nature*, 335:817–820, 1988.

[11] S. Peleg and J. Herman. Panoramic mosaics by manifold projection. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 338–343, 1997.

[12] L. Rabiner and B. H. Juang. *Fundamentals of Speech Recognition*. Prentice Hall, Englewood Cliffs, 1993.

[13] P. Rademacher and G. Bishop. Multiple-center-of-projection images. In *Proc. SIGGRAPH*, pages 199–206, 1998.

[14] S. M. Seitz and J. Kim. The space of all stereo images. *Int'l J. of Computer Vision*, 48:21–38, 2002.

[15] S. M. Seitz and J. Kim. Multiperspective imaging. *IEEE Computer Graphics and Applications*, 23:16–19, November/December 2003.

[16] H. Y. Shum and L. W. He. Rendering with concentric mosaics. In *Proc. SIGGRAPH*, pages 299–306, 1999.

[17] J. Stone. Object recognition using spatio-temporal signatures. *Vision Research*, 38(7):947–951, 1998.

[18] J. Stone. Object recognition: View-specificity and motion-specificity. *Vision Research*, 39(24):4032–4044, 1999.

[19] M. J. Tarr and H. H. Bülthoff. *Object recognition in man, monkey, and machine (cognition special issues)*. Cambridge, MIT Press, 1999.

[20] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 511–518, 2001.

[21] G. M. Wallis and H. H. Bülthoff. Effect of temporal association on recognition memory. In *Proc. Natl. Acad. Sci. USA*, volume 98, pages 4800–4804, 2001.

[22] G. Wolberg. *Digital Image Warping*. IEEE Computer Society Press, 1990.

[23] D. N. Wood, A. Finkelstein, J. F. Hughes, C. E. Thayer, and D. H. Salesin. Multiperspective panoramas for cel animation. In *Proc. SIGGRAPH*, pages 243–250, 1997.
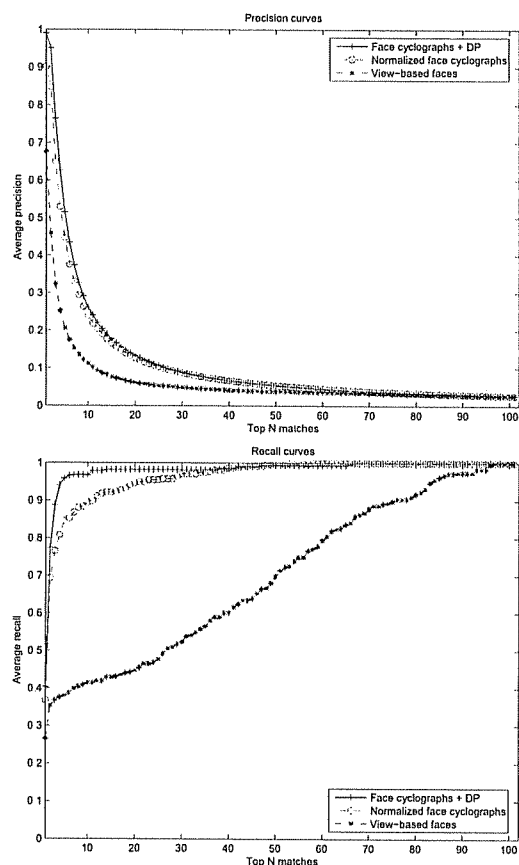
Figure 9: Average precision and recall with respect to the top N matches. The comparison is between the face cyclographs (multiperspective) and view-based faces (single perspective). In face cyclographs, two methods are proposed, one is face cyclographs + DP, and the other uses normalized face cyclographs.