

2.5D Elastic Graph Matching Algorithms

Stefanos Zafeiriou[†] Maria Petrou[†]
[†] Imperial College London
{s.zafeiriou,maria.petrou}@imperial.ac.uk

Vasileios Argyriou *
*Kingston University
{v.argyriou}@kingston.ac.uk

Abstract

In this paper, a series of advances in Elastic Graph Matching (EGM) for face recognition assisted by the availability of 3D facial geometry is proposed. More specifically, we conceptually extend the EGM algorithm in order to exploit the 3D nature of human facial geometry for face recognition/verification. In order to achieve that, first we extend the matching module of the EGM algorithm in order to capitalize on the 2.5D facial data. Moreover, we incorporate the 3D geometry into the multiscale analysis used and build a novel geodesic multiscale morphological pyramid of dilations/erosions in order to fill the graph jets. We demonstrate the efficiency of the proposed advances in the face recognition/verification problem.

1. Introduction

Dynamic link architecture (DLA) has been proposed as an abstract methodology for distortion invariant object recognition [9]. In DLA, an object is represented as a connected graph. Each graph node is located at a certain spatial image location \mathbf{x} . A feature vector, the so called jet, is attached at each graph node. The jet elements can be local brightness values that represent the image region around the node. However, it is desirable to have more complex types of jet that are produced by multiscale image analysis [9]. The representation of an object in DLA can be summarized in the following steps: 1) Group all the features that correspond to the same graph node of the object into a jet. 2) Group all the nodes and jets that belong to the object in order to form the object graph. 3) Define neighborhood relationships for each graph node.

One of the most well-studied implementations of DLA is the Elastic Graph Matching (EGM) algorithm. In [9] EGM was initially proposed for arbitrary object recognition from images and has been a very popular topic of research for various facial image characterization applications. In EGM, a reference object graph is created by overlaying a rectangular elastic sparse graph on the object image and then calculating a Gabor wavelet bank response at each graph node.

The graph matching process is implemented by a stochastic optimization of a cost function which takes into account both jet similarities and grid deformations. One of the basic advantages of EGM algorithms is that they can be applied in a fully automatic manner when combined with a face detector and/or a fully automatic facial feature detection algorithm [14, 11, 13].

A lot of research has been conducted in order to boost the performance of EGM for face recognition, face verification and facial expression recognition [5, 8, 14, 13, 11]. In [5], EGM has been proposed and tested for frontal face verification. A variant of the typical EGM, the so-called Morphological Elastic Graph Matching (MEGM), has been proposed for frontal face verification and tested under various recording conditions [8, 14, 13]. EGM with morphological feature vectors for facial expression recognition was proposed in [12].

In this paper, we present EGM methods that exploit the 3D facial geometry. Nowadays, the acquisition of 3D faces is a relatively simple procedure due to the low cost devices that are available. This has constituted 3D face recognition a very rapidly growing research topic. Another reason, that resulted in the increase of 3D face recognition popularity as a research topic, is that although 2D face recognition systems can achieve good performance in constrained environments, they still encounter difficulties in handling large amounts of facial variations due to head pose, lighting conditions and facial expressions. Moreover, the acquisition of 3D face geometry is also quite restricted due to the presence of noise and the difficulty presented in image acquisition [3]. For the acquisition of 3D data, different techniques have been used, such as laser and structured light scanners. However, the use of these techniques is not very practical as it requires the cooperation of the subject. So, we use photometric stereo [2] for 3D acquisition, that does not depend significantly on the user's cooperation. The actual model that we acquire this way is a 2.5D model. 2.5D is a simplified 3D (x, y, z) surface representation that contains at most one depth value (z direction) for every point in the (x, y) plane.

All EGM based methods proposed so far do not take into

consideration the 3D nature of the objects. In this paper we propose methods for exploiting the 3D nature of the objects inside the EGM architecture. That is, the proposed algorithm introduces a new structure for object representation combining both image intensity and 3D geometry information. Moreover, we propose algorithms which can be used for robustly matching the object structure in novel instances. The algorithm is general and can be used for arbitrary object recognition. In this paper we demonstrate the power of the proposed method in face recognition assisted by the available 3D face geometry. Moreover, we exploit the 3D face information available in order to advance the matching procedure. Moreover, we adopt the assumption that facial expressions are isometries upon the facial surface and we incorporate this inside the matching algorithm. This assumption has been exploited in [4, 10] for creating expression invariant face representations.

The proposed method in its generalized version can be applied in a fully automatic manner, combined with a face detection algorithm. The simplified version of the proposed algorithm requires the detection of the nose. The nose is the most easily detected facial feature when the 3D facial geometry is available. On the contrary, most of the 3D face recognition algorithms proposed so far require the robust detection [7] of a set of facial features, which is in general a difficult task. Furthermore, we incorporate the 3D facial geometry inside the multiscale analysis and we propose a novel Geodesic Morphological Multiscale analysis in order to fill the graph jets.

2. 2D Elastic Graph Matching Techniques

In the first step of the EGM algorithm, a suitable for face representation sparse graph is selected. An evenly distributed rectangular graph is one of the most easy to handle automatically forms of a graph, since only a face detection algorithm is needed in order to find an initial approximation of the rectangular facial region. The facial image region is analyzed and a set of local descriptors is extracted at each graph node (called jets). Analysis is usually performed by building an information pyramid using scale-space techniques. In the standard EGM, a 2D Gabor based filter bank has been used for image analysis. The output of multiscale morphological dilation-erosion operations or the morphological signal decomposition at several scales are nonlinear alternatives of the Gabor filters for multiscale analysis and both have been both successfully used for facial image analysis [8]. Morphological feature vectors are robust against plane rotations. A jet based on Gabor filters that is robust against rotation and scaling has been recently proposed in [11].

Formally, at each graph node that is located at image co-

ordinates \mathbf{x} , a jet (feature vector) $\mathbf{j}(\mathbf{x})$ is formed:

$$\mathbf{j}(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_M(\mathbf{x})]^T, \quad (1)$$

where $f_i(\mathbf{x})$ denotes the output of a local operator applied to image f at the i -th scale or at the i -th pair (scale, orientation) and M is the jet dimensionality.

The next step of EGM is to match the reference graph on the test facial expression image in order to find the correspondences of the reference graph nodes on the test image. This is accomplished by minimizing a cost function that employs node jet similarities while preserving at the same time the node neighborhood relationships. Let the subscripts t and r denote a test and a reference facial image (or graph), respectively. The L_2 norm between the feature vectors at the l -th graph node of the reference and the test graph is used as a similarity measure between jets. Another similarity measure that is used is the cosine between the two vectors:

$$C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l)) = \frac{\mathbf{j}(\mathbf{x}_r^l)^T \mathbf{j}(\mathbf{x}_t^l)}{\|\mathbf{j}(\mathbf{x}_r^l)\| \|\mathbf{j}(\mathbf{x}_t^l)\|}. \quad (2)$$

Let \mathcal{V} be the set of all graph vertices of a certain facial image. For the rectangular graphs, all nodes apart from the boundary nodes have exactly four connected nodes. Let $\mathcal{N}(l)$ be the four-connected neighborhood of node l . In order to quantify the node neighborhood relationships using a measure, the relative normalized local node deformation is used:

$$C_d(\mathbf{x}_t^l, \mathbf{x}_r^l) = \sum_{\xi \in \mathcal{N}(l)} \frac{\|(\mathbf{x}_t^l - \mathbf{x}_t^\xi) - (\mathbf{x}_r^l - \mathbf{x}_r^\xi)\|}{\|\mathbf{x}_r^l - \mathbf{x}_r^\xi\|}. \quad (3)$$

The set of vertices $\{\mathbf{x}_t^l(r)_{\text{opt}}\}_{l=1}^L$ can be found by minimizing:

$$\begin{aligned} \{\mathbf{x}_t^l(r)_{\text{opt}}\}_{l=1}^L &= \min_{\{\mathbf{x}_t^l\}} C(\{\mathbf{x}_t^l\}) \\ &= \min_{\{\mathbf{x}_t^l\}} \sum_{l \in \mathcal{V}} \{-C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l)) \\ &\quad + \lambda C_d(\mathbf{x}_t^l, \mathbf{x}_r^l)\}. \end{aligned} \quad (4)$$

The choice of λ in (10) controls the rigidity/plasticity of the graph [5], [8].

In [8], the optimization of (4) was implemented as a Simulated Annealing (SA) procedure, that imposes global translation at the graph and local node deformations. In [12], in order to deal with face translation, rotation and scaling, the following optimization problem has been formulated:

$$\{\{\delta_l\}_{l=1}^L, \mathbf{t}, \mathbf{A}\}_{\text{opt}} = \min_{\{\delta_l\}_{l=1}^L, \mathbf{t}, \mathbf{A}} C(\{\mathbf{A}\mathbf{x}_r^l + \mathbf{t} + \delta_l\}) \quad (5)$$

The above problem can be interpreted as follows: find a wrapping matrix \mathbf{A} (which in our case is a scaling and rotation matrix), a global translation \mathbf{t} and a set of local node perturbations $\{\delta_l\}_{l=1}^L$ such that the cost (4) is minimized.

The optimal set of test node coordinate vectors is given by $\mathbf{x}_t^l = \mathbf{A}\mathbf{x}_r^l + \mathbf{t} + \delta_l$. The above optimization problem is solved using simulated annealing, as well.

In [12] the above optimization problem has been solved following a SA approach where random translations, scalings and rotations were tested. In order to find the local node perturbation, a local optimization problem is solved for each one of the nodes, as:

$$\begin{aligned} \{\delta_l\}_{\text{opt}} = \min_{\delta_l} & -C_f(\mathbf{j}(\mathbf{x}_r^l), \mathbf{j}(\mathbf{A}\mathbf{x}_r^l + \mathbf{t} + \delta_l)) \text{ s.t.} \\ & C_d(\mathbf{x}_r^l, \mathbf{A}\mathbf{x}_r^l + \mathbf{t} + \delta_l) \leq \tau_{\text{max}}, \text{ and, } \|\delta_l\| \leq \delta_{\text{max}}. \end{aligned} \quad (6)$$

The above algorithm can be summarized as follows.

- Initialize the algorithm using an initial guess $\{\mathbf{x}_t^{(1)}\}_{l=1}^L$ (using the result of a face detector).
- For a number of steps
 - choose a global random similarity transformation (i.e., rotation, translation and scaling) in a certain range;
 - for all nodes solve the local constrained optimization problem (6);
 - check whether you accept the new solution $\{\mathbf{x}_t^{(k)}\}_{l=1}^L$.
- End

2.1. A Robust Method using Simulated Annealing with a Reliability Criterion for Nodes

Another EGM method for handling rotation and scaling using Gabor based features was proposed in [11]. Initially a coarse step is applied in order to find an initial graph position. Afterwards, the nodes are randomly visited in order to match them locally. A set of reliable nodes is then selected. The reliability of the nodes is tested as follows. Let i_1, i_2, i_3 be three linked nodes. A relative angle, around a test node, is calculated as:

$$\theta^t(i_1, i_2, i_3) = \cos^{-1} \left(\frac{(\mathbf{x}_t^{i_2} - \mathbf{x}_t^{i_1})(\mathbf{x}_t^{i_3} - \mathbf{x}_t^{i_1})}{\|\mathbf{x}_t^{i_2} - \mathbf{x}_t^{i_1}\| \|\mathbf{x}_t^{i_3} - \mathbf{x}_t^{i_1}\|} \right) \quad (7)$$

Then, a test node is determined as reliable if it satisfies the following criterion:

$$|\theta^t(i_1, i_2, i_3) - \theta^r(i_1, i_2, i_3)| < \tau_\theta. \quad (8)$$

Having found this set of reliable nodes, we extract a wrapping matrix \mathbf{A} and a translation vector \mathbf{t} by solving a simple optimization problem. The optimization problem is as follows:

$$(\mathbf{A}, \mathbf{t})_{\text{opt}} = \arg \min_{\mathbf{A}, \mathbf{t}} \sum_{k \in \mathcal{L}} \|\mathbf{x}_t^k - (\mathbf{A}\mathbf{x}_r^k + \mathbf{t})\|^2 \quad (9)$$

where \mathcal{L} is the set of reliable nodes and \mathbf{t} represents the translational vector between graphs. For a set of six parameters, at least 3 reliable nodes should be identified. After solving (9), the optimal wrapping matrix \mathbf{A}_{opt} and translation vector \mathbf{t}_{opt} are applied to all nodes and the whole procedure is repeated. A series of steps, including scaling, rotation, translation and local node perturbations are then calculated. The algorithm is summarized as follows.

- Initialize the algorithm using a coarse matching procedure $\{\mathbf{x}_t^{(1)}\}_{l=1}^L$.
- For a number of steps
 - for all nodes solve the local optimization problem (6);
 - if the reliable nodes (reliable according to (8)) are more than 3, then estimate the wrapping matrix \mathbf{A}_{opt} and the translation vector \mathbf{t}_{opt} ;
 - re-estimate the position of all nodes according to the wrapping matrix \mathbf{A}_{opt} and the translation vector \mathbf{t}_{opt} .
- If $\left| \frac{\sum_{l \in \mathcal{V}} \{-C_f(\mathbf{j}(\mathbf{x}_t^{(k)}), \mathbf{j}(\mathbf{x}_r^l))\}}{\sum_{l \in \mathcal{V}} \{-C_f(\mathbf{j}(\mathbf{x}_t^{(k-1)}), \mathbf{j}(\mathbf{x}_r^l))\}} \right| < \epsilon$, then stop.

In [11] the nodes were visited randomly and were locally optimized based on a SA strategy. Finally, the SA was cooled after the re-initialization of all nodes. The final measure $\sum_{l \in \mathcal{V}} \{-C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l))\}$ was used for quantifying the distance between the reference graph and the test image.

3. 2.5D Matching Algorithm

In this Section we describe the proposed matching algorithm which makes use of the available 2.5D facial information. The geodesic distances between the points can be calculated given the facial surface. Moreover, they are robust against isometry mappings (or isometries) of the facial surfaces. The facial pose variations are isometries of the facial surfaces. Up to a certain extent facial expressions can be safely regarded as isometries, as well, as long as the mouth remains closed. That is, the geodesic distances before and after the development of an expression are considered to remain (approximately) the same. Recently, a lot of research has been conducted in order to create expression invariant representation for face recognition [4, 10], where facial expressions have been modelled as isometries of the facial surface. The notion of geodesic distance, the geodesic path and the angle between geodesic paths are pictorially described in Figure 1 a.

By incorporating the isometry assumption in the EGM architecture, a matching procedure is implemented in which

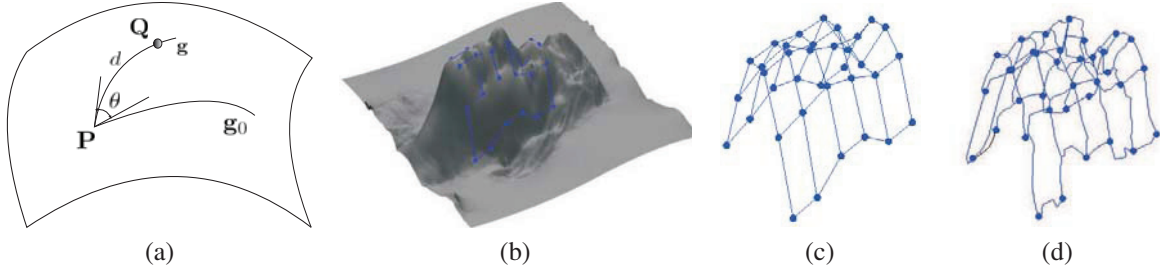


Figure 1. a) Geodesics paths and geodesic angles; b) the facial geometry and the graph; c) the 3D grid; d) the 3D geodesic graph.

the nodes can be deformed in such a way that the geodesic distances are approximately preserved during the matching procedure. In our case, the 2.5D information was derived from photometric stereo [2] and the integration of the normal field [6]. Thus a height map $g(\mathbf{x}) \in \mathbb{R}$ was acquired for every pixel of the image as $\mathbf{x} \in \mathbb{Z}^2$. Let \mathcal{X}_g be a set of connections \mathcal{X}_g between the points of the surface which defines the topology of the surface. The set of connections is derived from a triangulation of the image domain (in our implementation we used a simple Delaunay triangulation). Let that $\mathcal{G} = (g, \mathcal{X}_g)$ denote the tuple that contains both the depth map and the topology of the facial surface.

Optimization problem (4) was modified in such a way as to find a set of vertices $\{\mathbf{x}_t^l(r), l \in \mathcal{V}\}$ in the test image that minimizes the following cost function:

$$C_{\mathcal{G}}(\{\mathbf{x}_t^l\}) = \sum_{l \in \mathcal{V}} \left(-\{C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l)) + \lambda \sum_{\xi \in \mathcal{N}(l)} \left(\frac{|d_{\mathcal{G}}(\mathbf{x}_t^l, \mathbf{x}_r^{\xi}) - d_{\mathcal{G}}(\mathbf{x}_r^l, \mathbf{x}_r^{\xi})|}{d_{\mathcal{G}}(\mathbf{x}_r^l, \mathbf{x}_r^{\xi})} \right) \right) \quad (10)$$

where $d_{\mathcal{G}}(\mathbf{x}_1, \mathbf{x}_2)$ is the geodesic distance between points \mathbf{x}_1 and \mathbf{x}_2 of the triangulated surface \mathcal{G} . If we consider that the neighborhood of \mathcal{V}_i is sufficiently local (i.e., we have a dense graph and a topology without holes) then we may approximate:

$$d_{\mathcal{G}}(\mathbf{x}_1, \mathbf{x}_2) \approx \|\mathbf{z}(\mathbf{x}_1) - \mathbf{z}(\mathbf{x}_2)\| \quad (11)$$

where $\mathbf{z}(\mathbf{x}) = [\mathbf{x} \ g(\mathbf{x})]^T$.

The corresponding local optimization problem (5) in order to estimate the local perturbation is now formulated as:

$$\{\delta_l\}_{\text{opt}} = \min_{\delta_l} -C_f^2(\mathbf{j}(\mathbf{x}_r^l), \mathbf{j}(\mathbf{A}\mathbf{x}_r^l + \mathbf{t} + \delta_l)) \text{ s.t.} \\ d_{\mathcal{G}}(\mathbf{x}_r^l, \mathbf{A}\mathbf{x}_r^l + \mathbf{t} + \delta_l) \leq \tau, \text{ and, } \|\delta_l\| \leq \delta_{\text{max}}. \quad (12)$$

The local optimization problem is schematically described in Figure 4 a.

The proposed EGM algorithm that uses the available 3D facial geometry is as follows. First, a coarse step is applied and a first approximation of the graph is calculated. The coarse step can be a simple implementation of the step described in the previous Section. For every node a constrained local search is performed in order to solve (12). Moreover, instead of using \mathbf{x} , $\mathbf{z}(\mathbf{x})$ can be used so as to find

a 3×3 wrapping matrix \mathbf{A} . In this case at least 3 reliable nodes should be identified. After all nodes have been visited, the set of reliable nodes is extracted. In our case, since 2.5D information is available and in order to retain robustness against pose variations, criterion (8) is generalized as follows. Let i_1, i_2, i_3 be three linked nodes. The relative angle around a test node is calculated as:

$$\theta^t(i_1, i_2, i_3) = \cos^{-1}(\theta_{\mathcal{G}}(i_1, i_2, i_3)) \quad (13)$$

where $\theta_{\mathcal{G}}$ is the angle between the geodesic path i_1, i_2 and the geodesic path i_1, i_3 (as shown in Figure 2). In case the nodes are close enough, then:

$$\theta(i_1, i_2, i_3)^t = \cos^{-1}(\theta_{\mathcal{G}}(i_1, i_2, i_3)) \\ \approx \cos^{-1} \left(\frac{(\mathbf{z}(\mathbf{x}_r^{i_3}) - \mathbf{z}(\mathbf{x}_r^{i_1}))(\mathbf{z}(\mathbf{x}_r^{i_2}) - \mathbf{z}(\mathbf{x}_r^{i_1}))}{\|\mathbf{z}(\mathbf{x}_r^{i_3}) - \mathbf{z}(\mathbf{x}_r^{i_1})\| \|\mathbf{z}(\mathbf{x}_r^{i_2}) - \mathbf{z}(\mathbf{x}_r^{i_1})\|} \right). \quad (14)$$

In any case, all angles are robust against out of plane facial pose changes. Examples of the relative geodesic and grid angles are depicted in Figure 2 a and b.

After the set of reliable nodes are found, a 3×3 wrapping matrix is calculated by solving a similar optimization as (9). Afterwards this wrapping matrix is used for recalculating all node positions. The whole procedure is repeated in a similar way as the algorithm in the previous Section. A matching result is shown in Figures 3 a and b.

In order to simplify the procedure, we shall introduce a modification of the generalized algorithm proposed above. Assume that only certain facial features, like the nose, are available. The nose is one of the most easily detected features upon the facial surface. Given the position of the nose (points $\mathbf{x}_{\text{nose}}^l$ and $\mathbf{x}_{\text{nose}}^t$) an initial estimation for the position of the graph is provided so that we can search for small translations. Afterwards, under the assumption that facial expressions are isometries on the facial surface, we confine our local search, in the local matching procedure, to areas in which the geodesic distance between the position of the reference node and the position of the test node does not significantly change. Given the position of the node we can

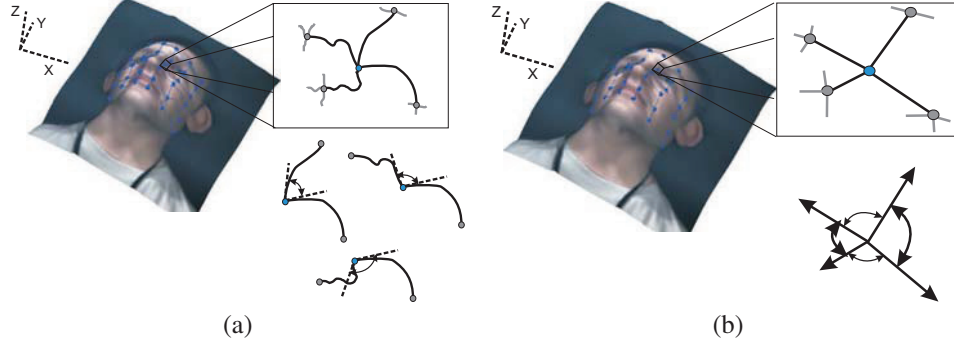


Figure 2. a) Relative geodesic angles in (13); b) relative angles in case (14)

modify the local search problem as:

$$\min_{\delta_l} \{-C_f(\mathbf{j}(\mathbf{x}_r^l + \delta_l), \mathbf{j}(\mathbf{x}_r^l))\} \text{ s.t.} \\ \left| \frac{d_G(\mathbf{x}_r^l, \mathbf{x}_r^l \text{ nose}) - d_G(\mathbf{x}_r^l + \delta_l, \mathbf{x}_r^l \text{ nose})}{d_G(\mathbf{x}_r^l, \mathbf{x}_r^l \text{ nose})} \right| \leq \tau_1, \quad \|\delta_l\| \leq \delta_{\max}. \quad (15)$$

The new local search is pictorially described in Figure 4 b.

4. Geodesic Multiscale Morphological Elastic Graph Matching

In the previous Section we described a method that incorporates the available 2.5D information in the matching procedure in order to obtain a pose invariant matching. Moreover, the matching is to some extent invariant to facial expressions, as well. In this Section, we present features that are robust against facial pose variations and facial expressions. To that end we describe a Multiscale Morphological Analysis (MMA) that is robust against pose variations, the so-called Geodesic MMA (GMMA). Moreover, under the isometry assumption (i.e., preservation of geodesic distances) the proposed MMA analysis can be until some extent robust under facial expression development.

MMA for facial image analysis has been initially proposed in [8] and has been well incorporated in a Morphological EGM (MEGM) architecture. Another modified MEGM has been presented in [13]. In MEGM an information pyramid is built using multiscale morphological dilations-erosions. Given an image $f(\mathbf{x}) : \mathcal{D} \subseteq \mathcal{Z}^2 \rightarrow \mathbb{R}$ and a structuring function $h(\mathbf{x}) : \mathcal{H} \subseteq \mathcal{Z}^2 \rightarrow \mathbb{R}$, the dilation of the image $f(\mathbf{x})$ by $h(\mathbf{x})$ is denoted by $(f \oplus h)(\mathbf{x})$. Its complementary operation, the erosion, is denoted by $(f \ominus h)(\mathbf{x})$ [8]. The multiscale dilation-erosion pyramid of the image $f(\mathbf{x})$ by $h_\sigma(\mathbf{x})$ is defined as:

$$(f * h_\sigma)(\mathbf{x}) = \begin{cases} (f \oplus h_\sigma)(\mathbf{x}) & \text{if } \sigma > 0 \\ f(\mathbf{x}) & \text{if } \sigma = 0 \\ (f \ominus h_{|\sigma|})(\mathbf{x}) & \text{if } \sigma < 0 \end{cases} \quad (16)$$

where σ denotes the scale parameter of the structuring function. In [8] it was shown that the choice of the structuring

function does not lead to statistically significant changes in the verification performance. However, it affects the computational complexity of feature calculation. The efficient computation of dilations and erosions is a factor of great importance in practical applications.

In this paper we use, for simplicity of computations, only flat structuring elements (i.e., $h_\sigma(\mathbf{v}) = 0, \forall \mathbf{v} \in \mathcal{H}_\sigma, \|\mathbf{v}\| \leq |\sigma|$). For this type of structuring elements, the dilation and the erosion can be written as:

$$(f \oplus h_\sigma)(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{H}_\sigma} f(\mathbf{x} - \mathbf{y}) \quad (17) \\ (f \ominus h_{|\sigma|})(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{H}_\sigma} f(\mathbf{x} - \mathbf{y})$$

For a flat structuring function, dilations can be efficiently computed by applying running min/max calculations in which the computation of one scale exploits the previous outcome. That is, scale-recursive computations that speed up considerably feature calculation can be used. For a flat structuring function, scale-recursive max computations (i.e., multiscale image dilations) are based on the observation that:

$$(f \oplus h_\sigma)(\mathbf{x}) = \max\{(f \oplus h_\sigma)(\mathbf{x}), \max_{\mathbf{y} \in \Delta\mathcal{H}(\sigma+1)} \{f(\mathbf{x} + \mathbf{y})\}, \max\{f(\mathbf{x} \pm \mathbf{s})\}\} \quad (18)$$

where $\mathbf{s} = [\sigma + 1 \ \sigma + 1]^T$, set $\Delta\mathcal{H}(\sigma_{i+1}) = \{\mathbf{y} = [y_1 \ y_2]^T \in \mathcal{Z}^2 : \|\mathbf{y}\|^2 > \sigma^2, \|\mathbf{y}\|^2 \leq (\sigma + 1)^2, |y_1| \leq \sigma, |y_2| \leq \sigma\}$. A similar multiscale analysis can be applied for the calculation of the erosion. In this paper we capitalize on the available 2.5D information and propose a novel multiscale analysis. First, we define the surface depended structuring flat elements around \mathbf{x} as:

$$g_{\sigma, \mathbf{x}}(\mathbf{v}) = 0, \forall \mathbf{v} \in \{\mathbf{y} \in \mathcal{Z}^2 : d_G(\mathbf{v}, \mathbf{x}) \leq |\sigma|\}. \quad (19)$$

The new multiscale morphological analysis given a height map g with a topology \mathcal{X}_g is as follows:

$$(f \oplus g_{\sigma_i})_g(\mathbf{x}) = \max\{(f \oplus g_{\sigma_{i-1}})_g(\mathbf{x}), \max_{(\mathbf{v} \in \Delta\mathcal{H}_g(\mathbf{x}, \sigma_i))} \{f(\mathbf{v})\}\} \quad (20)$$

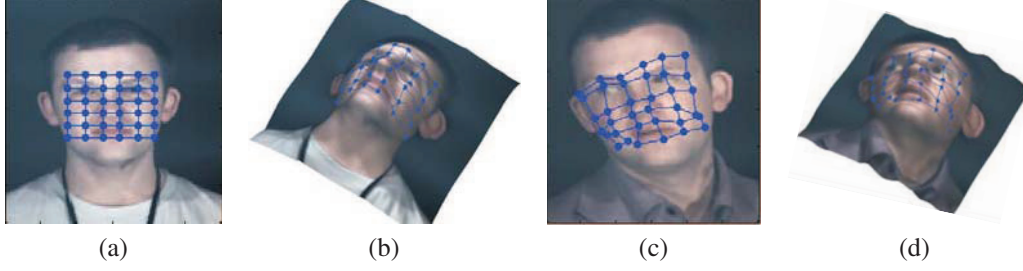


Figure 3. a) The 2D reference graph; b) the facial geometry and the graph; c) the matched graph on a test image; d) the facial geometry and the resultant graph.

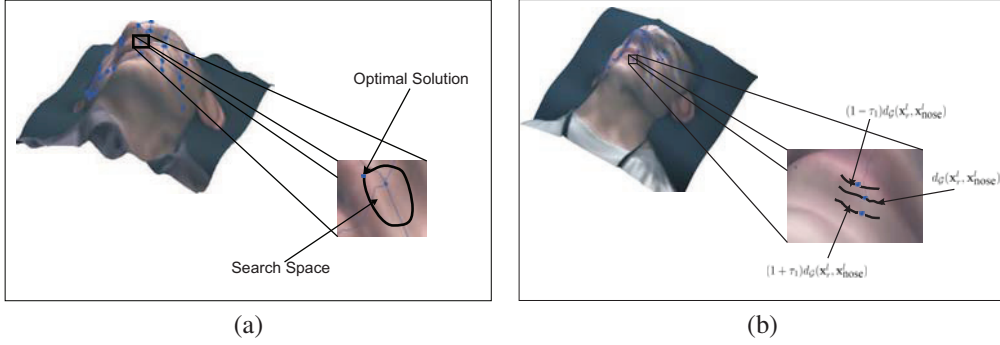


Figure 4. a) Local optimization problem (12); b) local optimization problem (15).

where $\Delta\mathcal{H}_G(\mathbf{v}, \sigma_i) = \{\mathbf{y} \in \mathcal{Z}^2 : \sigma_{i-1} < d_G(\mathbf{v}, \mathbf{y}) \leq \sigma_i\}$. In the same manner we have for the erosions:

$$(f \ominus g|_{\sigma_i})_G(\mathbf{x}) = \min_{\mathbf{v} \in \Delta\mathcal{H}_G(\mathbf{x}, \sigma_i)} \{f(\mathbf{v})\}. \quad (21)$$

In order to have a scale invariant GMMA we can modify the above as follows: $\Delta\mathcal{H}_G(\mathbf{v}, \sigma_i) = \{\mathbf{y} \in \mathcal{Z}^2 : (i-1)\sigma_1(\mathbf{v}, \mathbf{y}) < d_G(\mathbf{v}, \mathbf{y}) \leq i\sigma_1\}$ where $\sigma_1 = \mu \max\{d_G(\mathbf{x}_{\text{nose}}, \mathbf{x})\}$. $\max\{d_G(\mathbf{x}_{\text{nose}}, \mathbf{x})\}$ is the maximum geodesic distance between the nose and the given facial surface. In this paper we set $\mu = 1/(\max\{R_1, R_2\})$ where R_1 and R_2 are the number of image rows and columns, respectively.

The problem now becomes that of the efficient computation of GMFA in a computationally efficient manner. For the first scale, for all \mathbf{x} we have to compute the geodesic circle of radius σ_1 , i.e. $\mathcal{H}_{\sigma_1}(\mathbf{x}) = \{\mathbf{y} \in \mathcal{Z}^2 : d_G(\mathbf{x}, \mathbf{y}) \leq \sigma_1\}$. If σ_1 is small enough then we can substitute $d_G(\mathbf{x}, \mathbf{y}) \approx \|\mathbf{z}(\mathbf{x}) - \mathbf{z}(\mathbf{y})\|$. The calculation of $\mathcal{H}_{\sigma_2}(\mathbf{x})$ for the next scale can be performed by exploiting the property that if $\mathbf{y} \in \mathcal{H}_{\sigma_1}(\mathbf{x})$ then $\mathbf{x} \in \mathcal{H}_{\sigma_1}(\mathbf{y})$. This leads to $\sigma_2 = 2\sigma_1$:

$$\begin{aligned} \mathcal{H}_{2\sigma}(\mathbf{x}) &= \bigcup_{\mathbf{y} \in \mathcal{H}_{\sigma}(\mathbf{x})} \mathcal{H}_{\sigma}(\mathbf{y}) \\ \Delta\mathcal{H}_G(\mathbf{x}, 2\sigma) &= \mathcal{H}_{2\sigma}(\mathbf{x}) - \mathcal{H}_{\sigma}(\mathbf{x}). \end{aligned} \quad (22)$$

In the general case, we have:

$$\begin{aligned} \mathcal{H}_{\sigma_i}(\mathbf{x}) &= \bigcup_{\mathbf{y} \in \mathcal{H}_{\sigma_{i-1}}(\mathbf{x})} \mathcal{H}_{\sigma_{i-1}}(\mathbf{y}) \\ \Delta\mathcal{H}_G(\mathbf{x}, \sigma_i) &= \mathcal{H}_{\sigma_i}(\mathbf{x}) - \mathcal{H}_{\sigma_{i-1}}(\mathbf{x}) \end{aligned} \quad (23)$$

Thus, we only have to find a look-up table for the first scale and the next scales can be computed in a scale recursive strategy by using (23).

We can therefore build the GMMA $(f * h_{\sigma})_G(\mathbf{x})$ as (16). The output of these morphological operations forms the jet $\mathbf{j}(\mathbf{x}^l)$, at graph node l , that is located at image coordinates \mathbf{x}^l :

$$\begin{aligned} \mathbf{j}(\mathbf{x}^l) &= ((f * h_{\sigma_\Lambda})_G(\mathbf{x}^l), \dots, (f * h_{\sigma_1})_G(\mathbf{x}^l), f(\mathbf{x}^l), \\ &\quad (f * h_{\sigma_{-1}})_G(\mathbf{x}^l), \dots, (f * h_{\sigma_{-\Lambda}})_G(\mathbf{x}^l)) \end{aligned} \quad (24)$$

where Λ is the number of different scales used. An example of the typical multiscale morphological operations proposed in [8] for a flat structuring element is shown in Figure 6 a. The geodesic multiscale morphological analysis is shown in Figure 6 b.

5. Experimental Results

In this Section, we present experiments for face recognition using photometric stereo. A database of faces was collected setting a device for proper capture of four images under four different lighting directions. This is the first large photometric stereo face database that is collected and we are now try to make it publicly available. The four intensity images were processed using a standard photometric stereo method [1]. This resulted in a dense field of surface normals, which we integrated to form height maps using the well-known Frankot and Chellappa method [6].

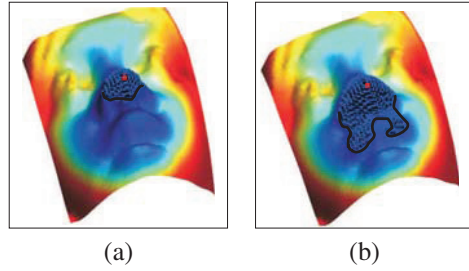


Figure 5. a) Geodesic circle of $2\sigma_1$; b) geodesic circle of $5\sigma_1$



Figure 6. a) Multiscale dilations-erosions for 9 scales; b) Geodesic multiscale dilations-erosions.

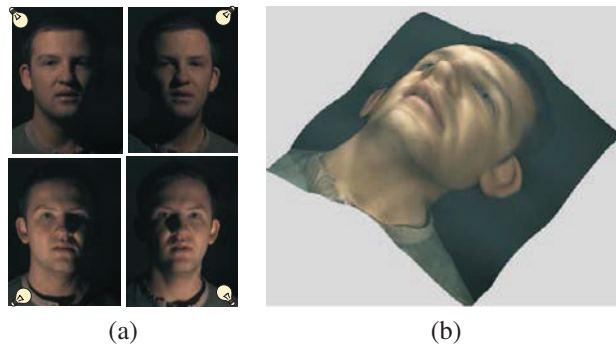


Figure 7. a) Four images taken for four different lighting directions; b) the corresponding 2.5D model.

The device was installed in the premises of General Dynamics. Staff and visitors were invited to use it. After a period of more than 6 months more than 250 persons were recorded. We had collected images for 113 persons, taken with more than a week's interval. For the majority of them (about 90 people) the interval was greater than one month. For the experiments presented here, a very challenging experimental procedure was followed using only one grayscale albedo image and the triangulated height map in order to create the reference graph and one grayscale albedo image and a triangulated height map for testing. We have also tested Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF). We use the abbreviation 2.5D-MEGM for MEGM using the 2.5D matching procedure presented in Section 3. The abbreviation GMEGM holds for the geodesic elastic graph matching with the 2.5D matching procedure. The recognition rates for the tested methods are summarized in Table 1.

Verification experiments have been conducted in the same database as well. In the test phase we have only one image for measuring the false rejection (FR) rates and the rest of them are used for measuring the false acceptance (FA) rates. The Receiver Operation Characteristics (ROC) curves are summarized in Figure 8. As it can be seen, the proposed methods advance the performance of EGM for face recognition/verification.

6. Conclusions

In this paper a series of advances in the EGM algorithm have been proposed in order to capitalize on the available 2.5D facial information using modern techniques. We have modified the matching technique in order to be robust against pose changes and also take into consideration changes of the facial surfaces under facial expressions. Moreover, we incorporated the 2.5D information in a multi-

Table 1. The recognition results of the tested methods

PCA	NNMF	MEGM	.2.5D-MEGM	GMEGM
46%	72%	78%	82%	87%

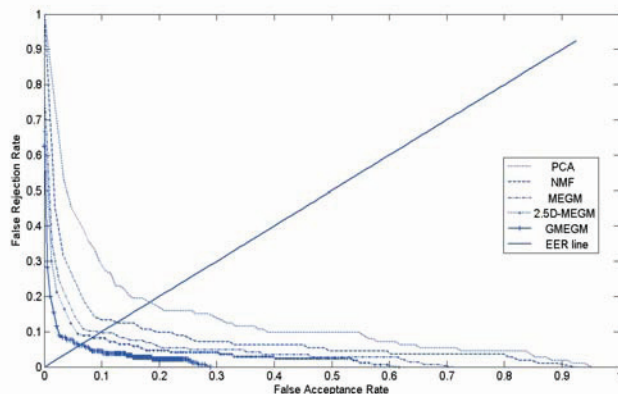


Figure 8. ROC curves for the tested methods.

scale morphological analysis and built that way a geodesic multiscale morphological analysis which is robust against facial pose changes. Experimental results show the proposed approaches advance the performance of EGM. We have collected the first of its kind photometric stereo face database of this size with more than 250 persons and we try to have it publicly available.

Acknowledgment

This work has been supported by the EPSRC project EP/E028659/1 Face Recognition using Photometric Stereo.

References

- [1] V. Argyriou and M. Petrou. Photometric stereo: An overview. *Advances in Imaging and Electronic Physics*, 156:1–55, 2008.
- [2] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (25):1239–1252, 2003.
- [3] K. Bowyer, K. Chang, and P. Flynn. A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition. *Computer Vision and Image Understanding*, 101(1):1–15, 2006.
- [4] A. Bronstein, M. Bronstein, and R. Kimmel. Expression-invariant representations of faces. *IEEE Transactions on Image Processing*, 16(1):188–197, 2007.
- [5] B. Duc, S. Fischer, and J. Bigün. Face authentication with Gabor information on deformable graphs. *IEEE Transactions on Image Processing*, 8(4):504–516, Apr. 1999.
- [6] R. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (10):439–451, 1988.
- [7] I. Kakadiaris, G. Passalis, G. Toderici, M. N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):640–640, 2007.
- [8] C. Kotropoulos, A. Tefas, and I. Pitas. Frontal face authentication using discriminating grids with morphological feature vectors. *IEEE Transactions on Multimedia*, 2(1):14–26, Mar. 2000.
- [9] M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. v. d. Malsburg, R. P. Würtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42(3):300–311, Mar. 1993.
- [10] I. Mpiperis, S. Malassiotis, and M. G. Strintzis. 3-d face recognition with the geodesic polar representation. *IEEE Transactions on Information Forensics and Security*, 2(3):537–547, 2007.
- [11] H.-C. Shin, J. Park, and S.-D. Kim. Combination of warping robust elastic graph matching and kernel-based projection discriminant analysis for face recognition. *IEEE Transactions on Multimedia*, 9(6):1125–1136, Oct. 2007.
- [12] S. Zafeiriou and I. Pitas. Discriminant graph structures for facial expression recognition. *IEEE Transactions on Multimedia*, 9(10):1528–1540, Dec. 2008.
- [13] S. Zafeiriou, A. Tefas, and I. Pitas. The discriminant elastic graph matching algorithm applied to frontal face verification. *Pattern Recognition*, 40(10):2798–2810, 2007.
- [14] S. Zafeiriou, A. Tefas, and I. Pitas. Learning discriminant person specific facial models using expandable graphs. *IEEE Transactions on Information Forensics and Security*, 2(1):50–55, Mar. 2007.