Correspondence

Gait Recognition Using Compact Feature Extraction Transforms and Depth Information

Dimosthenis Ioannidis, Dimitrios Tzovaras, Ioannis G. Damousis, Savvas Argyropoulos, and Konstantinos Moustakas

***Abstract*—This paper proposes an innovative gait identification and authentication method based on the use of novel 2-D and 3-D features. Depth-related data are assigned to the binary image silhouette sequences using two new transforms: the 3-D radial silhouette distribution transform and the 3-D geodesic silhouette distribution transform. Furthermore, the use of a genetic algorithm is presented for fusing information from different feature extractors. Specifically, three new feature extraction techniques are proposed: the two of them are based on the generalized radon transform, namely the radial integration transform and the circular integration transform, and the third is based on the weighted Krawtchouk moments. Extensive experiments carried out on USF "Gait Challenge" and proprietary HUMABIO gait database demonstrate the validity of the proposed scheme.**

***Index Terms*—Gait authentication, generalized radon transforms, genetic fusion, 3-D surface silhouette distribution.**

I. Introduction

Gait analysis has recently received growing interest within the computer vision community. Human movement analysis emerged a few decades ago mainly for medical analysis purposes [1], [2]. The latest research activities in multimodal biometrics evaluate the use of gait as a promising biometric modality. From a surveillance perspective, gait is an interesting modality because it can be acquired from a distance inconspicuously.

A. Current Approaches in Gait Recognition

Present work on automatic gait recognition has focused on the development of methods for extracting features from the input gait sequences. Gait analysis can be divided mainly into two techniques: model based and feature based (model free). Model-based approaches [3]–[7] study static and dynamic body parameters of the human locomotion. In [3], a multiview gait recognition method was presented using static activity-specific parameters, which are acquired from automatic segmentation of the body silhouette into regions. In [5], a gait recognition method has been proposed based on a statistical shape analysis. Moreover, in [7], a method was presented that extracts the gait signature from the evidence-gathering process. Experimental analysis in a dataset of ten subjects exhibited encouraging results. Conclusively, model-based approaches [3]–[7] create models of the human body from the input gait sequences. Previous work on these approaches shows that they are view and scale invariant. However, experimental evaluation in larger publicly available databases is needed in order to compare their performance to that of feature-based methods.

On the contrary, feature-based techniques used for gait recognition do not rely on the assumption of any specific model of the human body for gait analysis. Initially, the binary map of the moving person is estimated and a feature vector is extracted from the silhouette sequences. In [8], the extraction of features was performed on whole silhouettes; in [10], width vectors were used; in [11], Fourier descriptors were introduced; and, finally, in [12], angular transform was applied in silhouette sequences. All of the aforementioned techniques employ the use of gait in a temporal manner. The final stage in the feature-based approaches is the selection of the matching method to be used for finding the similarity between two input gait sequences. Proposed methods for matching are based on simple temporal correlation; full volumetric correlation on partitioned subsequent silhouette frames [8], [9]; linear time normalization [13]; and dynamic time warping [14], [15]. In most cases, Euclidean distance was used as a metric for distance calculation, but there are also reports on using procrustes distance [5] and symmetric group distances [16].

B. Motivation—The Proposed Approach

This paper proposes a novel gait identification and authentication method based on the use of novel 2-D and 3-D features of the image silhouette sequence. The proposed algorithm is tested and evaluated in two datasets and was compared to the state-of-the-art methods in gait analysis and recognition. Furthermore, the use of a genetic algorithm is proposed for fusing information from different feature extractors. Specifically, three new feature extraction techniques are proposed: two of them are based on the generalized radon transform, namely the radial integration transform (RIT) and the circular integration transform (CIT), which have been proven to offer a full analytical representation of the silhouette image using only a few descriptors, and the third is based on the weighted Krawtchouk moments that are well known for their compactness and discriminating power. The use of moments for shape recognition has recently received great attention [4], [16], [24]. Lee and Grimson computed a set of image features based on moments [4]. Shutler [24] proposed the use of the Zernike velocity moments for describing the motion throughout an image sequence. Zernike moments are based on a set of continuous orthogonal moment functions, such as Legendre moments. Experimental results on multiple datasets exhibited improvements in the recognition performance and illustrated their benefits over Cartesian velocity moments. One common problem with these moments is the discretization error, which increases as the order of the moment raises and, thus, limits the accuracy of the computed moments [25]. However, motivated by the successful use of these moments for gait recognition, this paper introduces the use of a set of discrete orthogonal moments, which do not involve any numerical approximation and are based on the weighted Krawtchouk polynomials [25], [26]. As a result, the error in the computed Krawtchouk moments is nonexistent and a reliable reconstruction of the original image can be achieved using relatively low-order moments. By using these weighted Krawtchouk moments, the recognition performance on the Gait-challenge database [9] will be seen to improve over the methods in [8], [13], [14], [32], and [33].

This paper also introduces the use of depth data, captured by a stereo camera for gait signal analysis. Depth-related data are assigned to the binary image silhouette sequences using two new transforms: the 3-D

Manuscript received October 31, 2006; revised May 28, 2007. This work was supported by the EU Cofunded Projects HUMABIO and STREP, 026990. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rama Chellappa.

The authors are with the Informatics and Telematics Institute, Thermi–Thessaloniki 57001, Greece (e-mail: djoannid@iti.gr; tzovaras@iti.gr; damousi@iti.gr; savvas@iti.gr; moustak@iti.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2007.902040

Radial silhouette distribution transform and the 3-D geodesic silhouette distribution transform. The proposed 3-D descriptors encode information not only about the position of the silhouette points on the image plane, but also about their 3-D distribution on the hull that represents the visible surface of the walking subject. The 3-D radial distribution transform encodes on each pixel of the extended silhouette image, the distance of the corresponding 3-D point from the center of mass of the 3-D hull. Even if this representation does not have any obvious physical meaning, it can be very easily extracted and is seen to provide satisfactory results. On the contrary, the 3-D geodesic silhouette transform provides very useful information about the degree of protrusion of each single point on the extended image. In particular, instead of measuring the distance of each point from the center of mass, the value stored on each point of the silhouette image corresponds to its degree of protrusion. The degree of protrusion is seen to implicitly capture information directly related to the underlying shape [19] (e.g., the final protrusion map of the silhouette will be different for subjects with larger main body or longer hands). Therefore, it is expected that this extended silhouette will be more suitable for the static identification of the shape of the walking subject.

The proposed algorithms are tested and evaluated in two large datasets and considerable improvements in recognition performance are seen in comparison to the state-of-the-art methods for gait recognition.

II. Preprocessing of Gait Sequences

A. Silhouette Extraction

In order to analyze the human movement, the walking subject silhouette needs to be extracted from the input image sequence. Initially, the background is estimated using a temporal median filter on the image sequence, assuming static background and moving foreground. Next, the binary silhouettes, denoted by B_k^{Sil}, are extracted by comparing each frame of the sequence with the background. The areas where the difference of their intensity from the background image is larger than a predefined threshold are considered silhouette areas. The generated silhouette images are noisy. Therefore, morphological filtering, based on antiextensive-connected operators [17], is applied so as to denoise the silhouette sequences. Finally, potential shadows are removed by analyzing the sequence in the HSV color space [18], thus resulting in the final binary silhouette $\tilde{B}_k^{\mathrm{Sil}}$.

B. Using Depth to Enhance Silhouette Sequence Extraction

A new technique is introduced that exploits depth data, if they are available. Let D_k denote the gait disparity data sequence and WC be a function that converts a disparity value $D(x,y)$ to 3-D coordinates in the world coordinate system using the already calibrated stereo-camera. At this stage of preprocessing, each gait sequence is composed of k binary silhouettes $\tilde{B}_k^{\mathrm{Sil}}(x,y)$, $x = 0,\ldots,R-1$, $y = 0,\ldots,C-1$, where R and C refer to the number of rows and columns of the binary silhouette, respectively. The world coordinates $(x_{WC,k}, y_{WC,k}, z_{WC,k})$ of the silhouette image of the kth frame are given by

$$(x_{WC,k}, y_{WC,k}, z_{WC,k}) = WC\,(x, y, D_k(x,y)) \cdot \tilde{B}_k^{\mathrm{Sil}}(x,y). \quad (1)$$

Subsequently, the center of mass $(x_{C,k}, y_{C,k}, z_{C,k})$ of each silhouette image is estimated.

In the following, the 3-D radial silhouette distribution and the 3-D geodesic silhouette distribution transform are defined. The proposed

(a) (b) (c) (d) (e)

Fig. 1. Illustration of silhouette representation utilized by the proposed system. (a) Binary silhouette, (b) 3-D radial distributed silhouette, (c) range map image, (d) silhouette triangulation, and (e) 3-D geodesic distributed silhouette.

3-D radial silhouette distribution transform, denoted as R_k^{Sil}, is defined as

$$R_k^{\mathrm{Sil}}(x,y) = \left\|WC\,(D_k(x,y)) - (x_{C,k}, y_{C,k}, z_{C,k})\right\|_2 \cdot \tilde{B}_k^{\mathrm{Sil}}(x,y) \quad (2)$$

$$\hat{R}_k^{\mathrm{Sil}}(x,y) = \frac{R_k^{\mathrm{Sil}}(x,y) - \min\left(R_k^{\mathrm{Sil}}(x,y)\right)}{\max\left(R_k^{\mathrm{Sil}}(x,y)\right) - \min\left(R_k^{\mathrm{Sil}}(x,y)\right)} \cdot 255 \quad (3)$$

where $\|\cdot\|_2$ denotes distance computation using L-2 Norm (Euclidean distance), k refers to the current frame of the gait silhouette sequence, (x,y) is the pixel position, and $\hat{R}_k^{\mathrm{Sil}}(x,y)$ represents the normalized 3-D silhouette values in the range from m to 255, as illustrated in Fig. 1(b). In the proposed scheme, m was selected to be equal to 60.

Instead of using Euclidean distance, the use of geodesic distance on the manifold of the silhouette surface is also proposed, with the introduction of the 3-D geodesic silhouette distribution transform. Initially, the triangulated version of the 3-D silhouette that also includes depth information is generated, as illustrated in Fig. 1(c) and (d). Adjacent pixels of the silhouette are grouped into triangles using Delaunay triangulation. Next, the dual graph $G = (V, E)$ of the given mesh is generated [20], where V and E are the dual vertices and edges. A dual vertex is the center of mass of a triangle and a dual edge links two adjacent triangles. The degree of protrusion for each dual vertex results from the following equation:

$$p(\mathbf{u}) = \sum_{i=1}^{N} g(\mathbf{u}, \mathbf{v}_i) \cdot \mathrm{area}(\mathbf{v}_i) \quad (4)$$

where $p(\mathbf{u})$ is the protrusion degree of dual vertex $\mathbf{u}$, $g(\mathbf{u}, \mathbf{v}_i)$ is the geodesic distance of $\mathbf{u}$ from dual vertex $\mathbf{v}_i$, and area $(\mathbf{v}_i)$ is the area of triangle that corresponds to the dual vertex $\mathbf{v}_i$.

Let us define $G_k'^{\mathrm{Sil}}(\mathbf{u})$ a function that refers to the dual vertices, to be given by

$$G_k'^{\mathrm{Sil}}(\mathbf{u}) = p(\mathbf{u}) \cdot \tilde{B}_k^{\mathrm{Sil}}(\mathbf{u}). \quad (5)$$

The 3-D geodesic silhouette distribution transform for the silhouette image, denoted as $G_k^{\mathrm{Sil}}(x,y)$, is simply a weighted average of the dual vertices that are adjacent to the corresponding pixel (x,y), i.e.,

$$G_k^{\mathrm{Sil}}(x,y) = \sum_{i=1}^{8} G_k'^{\mathrm{Sil}}(\mathbf{u}) \cdot w(x,y,\mathbf{u}) \quad (6)$$

$$\hat{G}_k^{\mathrm{Sil}}(x,y) = m + G_k^{\mathrm{Sil}}(x,y) \cdot (255 - m) \quad (7)$$

where $i = 1,\ldots 8$ denotes the number of adjacent pixels (x,y) to be weighted, $w(x,y,\mathbf{u})$ is the weighting function, and $\hat{G}_k^{\mathrm{Sil}}(x,y)$ represents the geodesic silhouette image at frame k, as illustrated in Fig. 2,

(a) (b)

Fig. 2. Illustration of 3-D geodesic transform. (a) Binary silhouettes and (b) corresponding 3-D geodesic distributed silhouettes.

with values in the interval of $[m, 255]$. In the proposed approach, m was selected to be equal to 60.

C. Normalization

In the final step of the preprocessing stage and before feature extraction, the preprocessed binary $(\hat{B}^{\mathrm{Sil}})$ or 3-D silhouette images $(\hat{R}^{\mathrm{Sil}}), \hat{G}^{\mathrm{Sil}})$ are scaled to the same resolution as in [9] and aligned to the center of the frame in each frame [9], [13], [14].

III. Feature Extraction From Gait Sequences

In this paper, the use of two 1-D radon transformations and the weighted Krawtchouk moments is proposed for generating the feature vector. In all cases, the input to the feature extraction system is assumed to be either the binary silhouettes $\left(\hat{B}_k^{\mathrm{Sil}}\right)$ or the 3-D–distributed silhouettes $\left(\hat{R}_k^{\mathrm{Sil}}, \hat{G}_k^{\mathrm{Sil}}\right)$ when the 3-D surface silhouette distribution transform is used.

A. 1-D Radon (RIT–CIT) Transformations

In this case, the feature extraction procedure transforms the input silhouette gait sequences using the RIT and the CIT [21]. Generalized radon transforms are proposed for feature extraction due to their capability to represent significant shape characteristics [22], [23]. Hence, it is expected that significant outcomes about a human shape and motion style can be exploited by studying these two generalized radon transforms. In particular, during human movement, there is a considerably large diversity in the angles of lower parts of the body (e.g., arms and legs), which should be unique among individuals. The proposed radon transforms ensure that these important dynamics of human shape will be encoded in the corresponding coefficients of RIT and CIT, thus enabling the classification of individuals using these compact feature transforms. Furthermore, another suitable recognition feature of the generalized radon transforms (RIT and CIT) is that each corresponding coefficient is estimated from the summation of several silhouette pixels and, thus, the proposed methods are less sensitive to the presence of noise on the silhouette image.

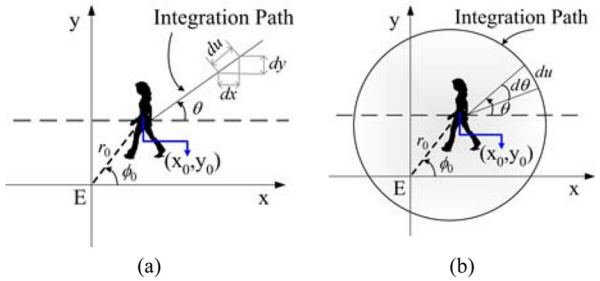
The RIT of a function $f(x, y)$ is defined as the integral of $f(x, y)$ in the direction of a straight line that starts from the point (x_0, y_o) and has angle θ with the horizontal axis x. The equation that calculates RIT for each θ is the following [21]:

$$\mathrm{RIT}_f(\theta) = \int_0^{+\infty} f(x_0 + u\cos\theta, y_0 + u\sin\theta)du \tag{8}$$

where u is the distance from the starting point (x_0, y_0).

In order to apply the RIT transform to the gait silhouettes, we assume that the origin is the center of mass (x_0, y_0) of the silhouette. Practically, since there in an infinite number of angles θ, the RIT transform is computed in steps of $\Delta\theta$. The angle step $\Delta\theta$ affects the level of detail of the transform. The discrete form of the RIT transform is used

$$\mathrm{RIT}(t\Delta\theta) = \frac{1}{J}\sum_{j=1}^{J}\mathrm{Sil}(x_0 + j\Delta u\cdot\cos(t\Delta\theta), y_0 + j\Delta u\cdot\sin(t\Delta\theta),)t = 1,\ldots,T \tag{9}$$



(a) (b)

Fig. 3. Applying the (a) RIT and (b) CIT transforms on a silhouette image using the center of gravity as its origin.

where Δu and $\Delta\theta$ are the constant step sizes of the distance (u) and angle (θ), J is the number of silhouette pixels that coincide with the line that has orientation θ and are positioned between the center of the silhouette and the end of the silhouette in that direction, Sil represents the correspondent binary or 3-D silhouette image, and, finally, $T = 360^\circ/\Delta\theta$ (Fig. 3).

Similarly, CIT is defined as the integral of a function $f(x, y)$ along a circle curve $h(\rho)$ with center (x_0, y_0) and radius ρ. The CIT is computed using the following equation:

$$\mathrm{CIT}_f(\rho) = \oint_{h(\rho)} f(x_0 + \rho\cos\theta, y_0 + \rho\sin\theta)du \tag{10}$$

where du is the arc length over the path of integration and $d\theta$ is the corresponding angle.

The center of the silhouette is again used as the origin for the CIT. The discrete form of the CIT transform is used, as depicted graphically in Fig. 3

$$\mathrm{CIT}(k\Delta\rho) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{Sil}(x_0 + k\Delta\rho\cdot\cos(t\Delta\theta), y_0 + k\Delta\rho\cdot\sin(t\Delta\theta) \tag{11}$$

where $k = 1,\ldots K$, $\Delta\rho$, and $\Delta\theta$ are the constant step sizes of the radius and angle variables, $K\Delta\rho$ is the radius of the smallest circle that encloses the binary or 3-D silhouette image Sil and, finally, $T = 360^\circ/\Delta\theta$.

Assume that a silhouette is scaled by λ in both directions. Then, the RIT and CIT of the scaled $\mathrm{Sil}_{SC}(x, y)$ is easily found to be [21]

$$\begin{aligned}\mathrm{RIT}_{\mathrm{Sil}_{SC}}(\theta) &= \lambda\cdot\mathrm{RIT}_{\mathrm{Sil}}(\theta),\\ \mathrm{CIT}_{\mathrm{Sil}_{SC}}(\rho) &= \lambda\cdot\mathrm{CIT}_{\mathrm{Sil}}\left(\frac{1}{\lambda}\rho\right).\end{aligned} \tag{12}$$

Otherwise stated, the RIT amplitude of the scaled silhouette image is only multiplied by the factor λ while the CIT of the scaled image is scaled by the λ factor and its amplitude is also multiplied by λ. Therefore, image scaling can affect the performance of the proposed method. For this reason, all gait sequences are normalized before feature extraction in order to overcome this scaling problem.

In addition, if the silhouette image function $\mathrm{Sil}(x, y)$ is written in polar form $\mathrm{Sil}(r, \phi)$ and $\mathrm{Sil}_{\mathrm{ROT}}(r, \phi) = \mathrm{Sil}(r, \phi - \phi_\alpha)$ is the rotated silhouette image by ϕ_α around the (r, ϕ) coordinate system's origin, then the RIT and the CIT of the rotated image are easily computed to be [21]

$$\begin{aligned}\mathrm{RIT}_{\mathrm{Sil}_{\mathrm{ROT}}}(\theta) &= \mathrm{RIT}_{\mathrm{Sil}}(\theta - \phi_\alpha)\\ \mathrm{CIT}_{\mathrm{Sil}_{\mathrm{ROT}}}(\rho) &= \mathrm{CIT}_{\mathrm{Sil}}(\rho)\end{aligned} \tag{13}$$

that is, the RIT of the rotated silhouette image is translated by ϕ_α and the CIT of the rotated image is unaffected.

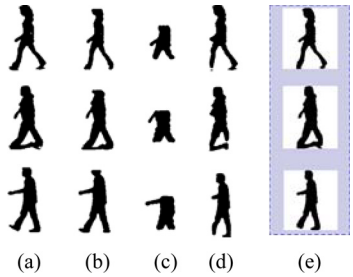


Fig. 4. Reconstruction of silhouette images using Krawtchouk moments for different moment order values (N, M). (a) Original silhouette $(W \times H = 188 \times 200)$, (b) $N = W/10, M = H/4$, (c) $N = W/10, M = H/16$, (d) $N = W/30, M = H/2$, and (e) $N = W/15, M = H/3$.

According to (13), when the input is rotated by ϕ_α, the RIT transform is also rotated with the same angle. This is a desirable feature in our case since the input gait sequences are captured in a near front-parallel view and the preprocessing step bounds the silhouette with a bounding box. Thus, in this case, the suggested Radon transform is not affected.

B. Feature Extraction Using Krawtchouk Moments

At almost all recent approaches on gait analysis, after feature extraction, the original gait sequence cannot be reconstructed. In the suggested approach, the use of a new set of orthogonal moments is proposed based on the discrete classical weighted Krawtchouk polynomials [25]. The orthogonality of the proposed moments ensures minimal information redundancy. In most cases, the Krawtchouk transform is used to extract local features of images [25]. The Krawtchouk moments Q_{nm} of order $(n + m)$ are computed using the weighted Krawtchouk polynomials for a silhouette image (binary or 3-D) with intensity function $\mathrm{Sil}(x, y)$ by [25]

$$Q_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \bar{K}_n(x; p1, N-1) \cdot \bar{K}_m(y; p2, M-1) \cdot \mathrm{Sil}(x, y) \tag{14}$$

$$\bar{K}_n(x; p, N) = K_n(x; p, N) \sqrt{\frac{w(x; p, N)}{\rho(n; p, N)}} \tag{15}$$

where $\bar{K}_n, \bar{K}_m$ are the weighted Krawtchouk polynomials, and $(N - 1) \times (M - 1)$ represents the pixel size of the silhouette image. Fig. 4 shows a graphical representation of the reconstructed silhouette images using different orders of N (for width) and M (for height).

Krawtchouk moments can be used to extract local information of the images by varying the parameters N and M. Parameter N can be used to increase the extraction of silhouette image in the horizontal axis. Larger N provides more information on the silhouette image in the horizontal axis, whereas the parameter M extracts local information of the silhouette image in the vertical axis. For the experiments, values for $N = R/15$ and $M = C/3$ were used, where R and C denote the number of rows and columns of the silhouette image, respectively. With these parameter values, a satisfactory reconstruction of the initial silhouette image can be achieved as illustrated in Fig. 4(e). Krawtchouk transform is proposed for feature extraction, due to its high discriminative power [26]. The benefits of using orthogonal moments for gait recognition has also been presented by Shutler [24], where Zernike moments were exploited to describe the shape and motion dynamics of gait sequences.

In our approach, weighted Krawtchouk moments are proposed for encoding the shape characteristics of silhouette sequences, which overcome the problems associated with continuous orthogonal moments, such as discretization errors and the need to transform coordinate spaces. The Krawtchouk transform is scale and rotation dependent. To remedy this problem, in our approach, the silhouette sequences are prescaled and aligned to the center; thus, the Krawtchouk transform is unaffected by scaling. However, a new set of Krawtchouk moments may also be used, which are rotation, scale, and translation invariant as presented in [25]. Finally, the input gait sequences are captured in a near fronto-parallel view and, thus, rotation does not affect the results of the Krawtchouk transform.

IV. Signature Matching Based on Weighted Classifiers Using a Genetic Algorithm

The following notations are used in this section: the term "gallery" is used to refer to the set of reference sequences, whereas the test or unknown sequences to be verified or identified are called "probe sequence." An important step in the recognition system, formally before the feature extraction stage, is gait cycle [12] detection of the gallery/probe sequence. In the literature [9], [13], [27], the detection of periodicity in a sequence is performed using a temporal analysis of the sequence. Temporal analysis can be applied on a set of different measurements, such as width of the bounding box, sum of the foreground pixels, etc. In this paper, a similar approach with [14] was followed, using autocorrelation of the input periodic signal.

A. Template-Matching Approaches

The methods that were exploited for gait template matching are based on the spatial–temporal correlation [9], [12] and dynamic time warping [14], [15]. Let $\mathbf{F}_{P,T}$ and $\mathbf{F}_{G,T}$ represent the feature vectors of the probe with N_P frames and the gallery sequence with N_G frames, respectively, and T denote a specific proposed transform (RIT, CIT, or KRAWTCHOUK).

In the spatial-temporal correlation method, the probe sequence is partitioned into consecutive subsequences of T_P adjacent frames, where T_P is the estimated period of the probe sequence. Also, let the kth probe subsequence be denoted as $\mathbf{F}^k_{(P,T)} = \left\{\mathbf{F}^{kT_P}_{P,T}, \ldots, \mathbf{F}^{(k+1)T_P}_{P,T}\right\}$ and the gallery sequence of N_G frames be denoted as $\mathbf{F}_{G,T} = \left\{\mathbf{F}^1_{G,T}, \ldots, \mathbf{F}^{N_G}_{G,T}\right\}$. Then, the distance metric between the kth subsequence, and the gallery sequence, for a specific feature transform T is defined as

$$\mathrm{Dist}_T(k) = \min_l \sum_{i=0}^{T_P-1} \sqrt{\sum_{x=0}^{S-1} \left(\mathbf{F}^{i+k\cdot T_P}_{P,T}(x) - \mathbf{F}^{i+l}_{G,T}(x)\right)^2} \tag{16}$$

where $k = 0, \ldots, m-1$ and $l = 0, \ldots, N_G - 1$ S denotes the size of a probe/gallery feature vector $\mathbf{F}$ for a specific transform, and $m = N_P/T_P$ represents the number of probe subsequences.

For robustness, after computing all distances between probe segments and gallery sequences of feature vectors, the median [9], [12] of the distances is taken as the final distance $D_T(\mathrm{Probe}, \mathrm{Gallery})$ between the probe and the gallery sequence

$$D_T = \mathrm{Median}\left(\mathrm{Dist}_T(1), \ldots, \mathrm{Dist}_T(m)\right), \quad m = \frac{N_P}{\mathrm{T}_P} \tag{17}$$

where m denotes the number of distances calculated between the probe subsequences and the whole gallery sequence. In (17), a smaller distance means a closer match between the probe and the gallery sequence.

Equation (16) indirectly assumes that the probe and gallery sequences are aligned in phase. Alternatively, in order to deal with the different walking speeds and intravariations of the gait cycles lengths, a dynamic time warping technique [15] is also examined. Specifically, the final distance D_T between a probe and a gallery sequence is calculated from the estimation of the distances between gait cycles, using the nonlinear rule as proposed in [14]. Experimental evaluation of this scheme in USF and proprietary HUMABIO databases exhibits slight

improvements in the recognition rates (varied from 0 to 1.5%) when compared to the spatial–temporal correlation approach. This slight increase was expected since the proposed algorithms are evaluated in two datasets, in which people are walking at similar speeds.

B. Genetic Algorithm (RCK-G) for Optimal RIT, CIT, and Krawtchouk Weights Calculation

In this section, a biometric fusion method at score level is proposed based on a genetic algorithm. Specifically, an optimum linear combination of matcher scores (distances) is exploited in order to improve the accuracy of the proposed scheme. A linear method based on a genetic algorithm is proposed instead of a typical Bayesian classifier due to the lack of the prior knowledge of the distribution of the estimated distances. Furthermore, genetic algorithms are very efficient optimization methods since they are capable of detecting near global optimum solutions without the need of *a priori* knowledge of the premise space and of any nonconvexities within it. They are also very scalable in terms of the number of variables to be optimized and can also be adapted to several problems with minimal effort. In the following paragraphs, the genetic algorithm [28], [29] that is proposed for the optimal weighting of the feature transforms is presented.

Let w_{RIT}, w_{CIT}, and w_{KR} be the weights for the RIT, CIT, and Krawtchouk, respectively. The w_{RIT}, w_{CIT}, and w_{KR} weights are concatenated in order to form the genotype or chromosome and range between 0 and 1. The training patterns (distance scores of the feature transforms) are also normalized within $[0, 1]$. Initially, a population of m chromosomes is generated. The quality of a specific chromosome is measured by calculating its fitness. The chromosome is decomposed to the weight values set $[w_{\mathrm{RIT}}, w_{\mathrm{CIT}}, w_{\mathrm{KR}}]$. Given this set, the similarity of each person (gallery) in the database to the client (probe) is calculated as follows:

$$\mathrm{Sim}(i,j) = \frac{w_{\mathrm{RIT}}}{D_{\mathrm{RIT}}(i,j)} + \frac{w_{\mathrm{CIT}}}{D_{\mathrm{CIT}}(i,j)} + \frac{w_{\mathrm{KR}}}{D_{\mathrm{KR}}(i,j)} \tag{18}$$

where i denotes the probe id, $i = 1, \ldots, NP$ (number of probes to identify), j denotes all subjects in the database, $j = 1, \ldots, NG$ and $D_T(i,j)$ is the final distance value (see (20), between the probe i and the gallery j for feature T, where $T = \{\mathrm{RIT}, \mathrm{CIT}, \mathrm{Krawtchouk}\}$ correspondingly. Then, the subject id C for which $\mathrm{Sim}(i,C) = \max(\mathrm{Sim}(i,j))$ is detected and a correct identification takes place if $C = i$. The fitness function for the chromosomes is computed as follows:

$$\begin{aligned}&\text{FitnessFunction}\\ &\quad = \sum_{i=1}^{NP} \text{correct\_id}_i\\ &\text{correct\_id}_i\\ &\quad = \begin{cases} 1, & \text{if } \mathrm{Sim}(i,C) = \max(\mathrm{Sim}(i,j)),\ j = 1, \ldots, NG\\ 0, & \text{if } \mathrm{Sim}(i,C) < \max(\mathrm{Sim}(i,j)),\ j = 1, \ldots, NG. \end{cases}\end{aligned} \tag{19}$$

It is obvious that as the fitness maximizes through the evolution of the population, so does the number of correctly identified individuals in the database. The GA described in [29] is allowed to run for 400 generations and the training on the proprietary database takes around 10 s for a population of 40 chromosomes (Pentium IV, 3.0-GHz CPU). The optimal weight set $[w_{\mathrm{RIT}}, w_{\mathrm{CIT}}, w_{\mathrm{KR}}]$ corresponds to the elite solution obtained by the GA at the final generation.

The GA-based weight optimization was performed using data not included in the gallery and probe sets of the experiments of Section VI in terms of different gait sessions. Specifically, the training on the USF dataset was performed using 30 subsets (each subset from a different subject) chosen randomly from the left and right camera of the near-side view (instead of using the frontside view). Similarly, the detection of optimal weights on the proprietary HUMABIO database was performed using 20 subsets from the normal and hat covariates of the available second session of the database (while for evaluation, the first session was used). For both databases, the optimal weights for the RIT transform vary from 40% to 42%; for CIT, they vary from 8% to 10%; and Krawtchouk weights vary from 49% to 51%. For all of the experimental tests, the following optimal weighted values were chosen:

$$w_{\mathrm{RIT}} = 0.4, \quad w_{\mathrm{CIT}} = 0.1, \quad w_{\mathrm{KR}} = 0.5.$$

The final weighted distance $D^W(i,j)$ between the current probe i and the gallery j is expressed as

$$D^W(i,j) = \frac{1}{\mathrm{Sim}(i,j)} \tag{20}$$

where $\mathrm{Sim}(i,j)$ denotes the maximum similarity between probe i and gallery j and is calculated using (17). In (20), a smaller weighted distance corresponds to a closer match between the probe and the gallery sequence.

Note that the proposed fusion method is only used to estimate the optimal weights. After the calculation of these weights, the trained algorithm is applied as is for the online identification of individuals and no further training or altering of the weights occurs. It does not refer to a training stage where the gallery set is used to estimate the weights that are optimally fitted to its content. This procedure might increase the recognition rates, but it would be very dependent to the training dataset and, thus, the effect of overtraining might appear for some of the probe sequences (especially for probe sequences captured under different conditions). Hence, here we only introduce a fusion at the score level whereas leaving our feature extraction algorithms to execute without any additional training procedures.

The proposed method was also compared to the simple min Euclidean distance on the matching scores. Using the latter method, a degradation of performance of about 7% was observed, which confirms the validity of our approach.

V. Databases' Description

The proposed methods were evaluated on two different databases: 1) the publicly available HumanID "Gait Challenge" dataset and 2) the proprietary large indoor HUMABIO dataset. The former database (described in detail in [9]) was captured in an outdoor environment and consists of people walking in elliptical paths in front of the camera. It is presently the largest available outdoor dataset and for each subject, image sequences were captured using up to five acquisition conditions: two camera angles (L, R), two shoe types (A, B), two surfaces (grass-G and concrete-C), with or without carrying a briefcase (BF), and time and clothing differences (N2). There are two available versions of the database. We have evaluated our algorithm on both versions. The first dataset consists of 71 subjects and can be evaluated using the seven experiments A-G. The latest version (v2.1) consists of 122 subjects and there are 12 experiments A-L. In the performed experiments, the GAR sequences are used as gallery and the probe sets contain individuals that are unique and there are no common sequences with the gallery sets.

The proprietary HUMABIO database was captured in an indoor environment. Currently, there are two available sessions that were captured with six months difference. The first session consists of 75 subjects and the second is 51, in which 48 subjects are common. Briefly, the collection protocol had each person walk multiple times naturally along a predefined path, so that the view is approximately fronto-parallel. The main course of walking is around 6 m on a concrete (C) surface. For each subject, up to three different conditions were captured: two shoe types (classic-CL and slipper-PA), with or without a hat (H)

TABLE I
COMPARISON OF THE PROPOSED METHODS TO THE BASELINE ALGORITHM [8] FOR ALL EXPERIMENTS A-G USING THE 1.7 VERSION OF THE USF DATABASE. (WE REMOVED FROM PROBE SETS D, E, F, G THE SUBJECTS THAT WERE NOT IN THE GALLERY)

Probe	Rank 1					Rank 1-5				
	RIT	CIT	KR	RCK-G	BASE [8]	RIT	CIT	KR	RCK-G	BASE
A (GAL) [71]	94	75	96	**97**	79	100	98	100	**100**	96
B (GBR) [41]	85	63	86	**89**	66	**95**	90	94	94	81
C (GBL) [41]	70	53	77	**83**	56	94	86	90	**94**	76
D (CAR) [66]	36	23	33	**41**	29	77	42	73	**78**	61
E (CBR) [42]	**36**	14	29	34	24	73	37	70	**72**	55
F (CAL) [66]	**31**	7	23	30	30	**60**	31	53	58	46
G (CBL) [42]	25	12	25	**28**	10	**63**	30	51	59	33

TABLE II
RECOGNITION PERFORMANCE OF THE PROPOSED WEIGHTED FEATURES (RCK–G) IN COMPARISON WITH OTHER APPROACHES FOR ALL EXPERIMENTS A-G ON THE 1.7 VERSION OF THE USF DATABASE

	Rank 1							Rank 1-5						
Exp		Similar Approaches not using training				Using training			Similar Approaches not using training				Using training	
	RCK-G	CMU [33]	LTN-S [13]	LTN-A [13]	DTW [14]	UMD-HMM [10]	MIT [32]	RCK-G	CMU	LTN-S	LTN-A	DTW	UMD -HMM	MIT
A	97	87	94	89	85	**99**	87	**100**	100	99	99	97	100	96
B	**89**	81	83	71	76	**89**	83	**94**	90	85	81	88	90	90
C	**83**	66	78	56	61	78	66	**94**	83	83	78	81	90	88
D	**41**	21	33	21	36	36	25	**78**	59	65	50	56	65	50
E	**34**	19	24	26	29	29	23	**72**	50	67	57	67	65	48
F	**30**	27	17	15	21	24	20	58	53	58	35	52	**60**	49
G	**28**	23	21	10	24	18	19	**59**	43	48	33	48	50	50

and (BF) when the subject carries a briefcase. For each sequence of each individual, the color (left and right) images were captured using the Bumblebee stereo camera from Point Grey Research. The camera lenses are precalibrated against distortion and misalignment and additional left and right images are aligned to within 0.05 pixel root mean square (rms) error. Furthermore, an optimized sum of squared differences approach (SSD) is utilized by the proposed scheme in order to estimate the disparity images [30]. This is the first database that has available accurate depth data for assisted gait recognition. In this paper, three experiments on this database are demonstrated (first session). Experiment A examines the hat covariate, B the briefcase, and, finally, C the shoe difference. In all cases, the gallery consists of classic shoe type (CL) without briefcase (NB) and not wearing a hat.

VI. EXPERIMENTAL RESULTS

The proposed framework was evaluated in the context of the gait identification and verification scenarios. Since the dynamic time warping approach exhibited overall improved recognition rates, the following results are presented using this technique for signature matching.

In some approaches for gait recognition, one subset of the gait sequences can be used for training such as in [31] and [32]. This approach increases the performance for some of the challenging experiments (time and clothing), but leads, in general, into a more complicated system in terms of building the trained gallery model. Moreover, these methods impose further constraints in the enrolment procedure, which should be significantly complicated. The proposed algorithm is extensively compared to methods that do not demand a training procedure. Furthermore, it is also compared with two state-of-the-art approaches that require training in a preprocessing stage.

For evaluation of the proposed approach in an identification scenario, cumulative match scores (CMS) are reported at ranks 1 and 5. Rank 1 performance illustrates the probability of correctly identifying subjects in the first place of the ranking score list and rank 5 illustrates the percentage of correctly identifying subjects in one of the first five places of the ranking score list.

A. Evaluation of the Proposed Gait System in the USF Database

The experimental results are reported in comparison with the USF baseline algorithm [9] for both versions (1.7 and 2.1) of the USF database. Moreover, comparative results are exhibited with other state-of-the-art gait recognition approaches on version 1.7 of the USF database. In particular, Table I compares the recognition performance of the proposed methods with the USF baseline algorithm [8]. It can be seen that the proposed feature classifiers give better rank-1 results on most experiments, thus validating the capability of the proposed transforms to encode the shape characteristics of the silhouette sequences by using relatively low-feature vector sizes. Furthermore, the suggested weighted combination of the feature descriptors using the genetic algorithm (RCK-G) outperforms the baseline algorithm in almost every experiment. In Table II, the recognition performance of the weighted features (RCK-G) method is presented in comparison with other known methodologies.

In particular, the results of RCK-G are compared with the method introduced in [33] that uses for identification features extracted from body and shape, the linear time normalization algorithm using the silhouette feature (LTN-S) [13] as well as using the angular transform (LTN-A) [13], the angular gait transform analysis using dynamic time warping [14], and finally for two methods from MIT [32] and UMD [10] that use training for building the gallery set, based on hidden Markov models. Table II illustrates that the proposed weighted feature algorithm outperforms all experiments with similar approaches that do not use training for building the gallery set. Furthermore, the proposed weighted algorithm (RCK-G) also achieves better performance for experiments C-G for rank 1 and 5 from the approaches that are based on training in a preprocessing state.

Moreover, the proposed methods were evaluated in the 2.1 version of the USF database that includes 122 subjects. The recognition results are presented in Table III in comparison with the USF baseline algorithm [9]. As seen, the proposed feature classifiers RIT and Krawtchouk give better rank-1 results on most experiments. Additionally, in challenging experiments when time and clothing (K, L) are appeared as a condition, RIT, CIT, and Krawtchouk algorithms perform better than the baseline

TABLE III
COMPARISON OF THE PROPOSED METHODS TO THE BASELINE ALGORITHM [9] FOR ALL EXPERIMENTS A-L USING THE 2.1 VERSION OF THE USF DATABASE

	Rank 1					Rank 1-5				
	RIT	CIT	KR	RCK-G	USF [9]	RIT	CIT	KR	RCK-G	USF
A	78	67	83	**83**	73	93	89	97	**96**	88
B	77	73	82	**86**	78	95	87	92	**94**	93
C	66	57	72	**78**	48	87	79	83	**88**	78
D	35	26	30	**39**	32	65	52	63	**66**	66
E	33	24	28	**34**	22	64	49	55	**63**	55
F	**22**	14	17	20	17	51	35	46	**51**	42
G	19	15	19	**21**	17	45	36	45	**46**	38
H	37	33	40	43	**61**	67	55	62	66	**85**
I	34	34	39	40	**57**	65	59	67	68	**78**
J	24	22	38	40	**36**	55	56	**68**	65	62
K	16	**17**	10	16	3	**54**	35	30	44	12
L	5	**6**	4	5	3	7	**28**	8	22	15

TABLE IV
RECOGNITION PERFORMANCE OF THE PROPOSED GENERALIZED RADON TRANSFORMS, KRAWTCHOUK AND THE RCK–G ON BINARY, RADIALLY, AND GEODESIC DISTRIBUTED SILHOUETTES

Probe Set	Rank 1 (%)					
Gallery C-CL-NB	RIT	CIT	Krawtchouk	RCK-G (binary)	RCK-G (radial)	RCK-G (geodesic)
A (Hat)	95	81	89	96	96	97
B (BF)	55	22	66	66	69	71
C (PA)	78	75	77	81	83	87
	Rank 1-5 (%)					
	RIT	CIT	Krawtchouk	RCK-G (binary)	RCK-G (radial)	RCK-G (geodesic)
	99	92	95	99	99	99
	71	43	79	79	80	84
	92	86	87	90	92	97

algorithm but the identification rate at rank 1 is still not very high. Finally, the proposed optimal weighted feature method using a genetic system (RCK-G) outperforms the baseline algorithm in almost every experiment.

B. Evaluation on the Proprietary HUMABIO Gait Database

The proposed methods were also evaluated using the proprietary HUMABIO gait database that includes gait data from 75 subjects. The experimental results are presented in Table IV for all proposed methods on all silhouette transforms. The proposed methodologies perform well in the experiments with the hat and the shoe (experiments A and C). In the briefcase experiment (experiment B), the recognition rates are lower. This confirms the results on the USF database and verifies the hypotheses that the briefcase experiment still remains a challenge. From the proposed feature extractors, RIT and Krawtchouk achieve, in general, better performance than the CIT. The proposed weighted feature methodology (RCK-G) was evaluated on the silhouette extracted features and the identification rate of the system was increased by 1% in hat (A) and 3% in slipper (C) experiment. Finally, the proposed scheme that utilizes 3-D silhouette maps was evaluated for the HUMABIO gait database.

It should be noted that the HUMABIO database includes depth data for the gait sequence. The proposed weighted feature algorithm based on 3-D silhouettes (radial and geodesic) achieve better performance on the experiments with briefcase and shoe (experiments B and C). Specifically, using the radial silhouette distribution transform, identification rates are increased by 3% for the briefcase (B) and 2% for the shoe (C) experiment, when compared to the case of the binary silhouette images. In addition, using the geodesic silhouette distribution transform, rank 1 is increased by 2% for the briefcase and 4% for the shoe experiment, when compared to the radial silhouette distribution transform. The briefcase experiment still remains a challenge, but the rank-1 rate has increased 5% when compared to the performance using binary silhouette sequences. The 3-D extended silhouettes perform up to 7% better when compared to the binary silhouettes. This result was expected during the design of the framework since 3-D information about the silhouette is present in both extended silhouettes. The further increased performance of the geodesic silhouette is also expected, since it encodes very useful information regarding the topology of the manifold that represents the 3-D silhouette.

(a) (b)

Fig. 5. (a) Original silhouettes. (b) Noisy silhouettes with $\sigma = 1.5$.

Finally, in order to evaluate the resilience of the proposed features, the extracted silhouettes were distorted with additive noise, as illustrated in Fig. 5. Specifically, the contour of the binary silhouettes was initially partitioned into small line segments. Then, Gaussian noise $N(0, \sigma)$ was added in the perpendicular direction of the line segments. The experimental evaluation did not exhibit a significant decrease in the recognition rates for small values of the standard deviation σ. The robustness of the proposed radon transforms and the Krawtchouk moments to additive noise were also proven in [21] and [25], respectively.

Besides identification, the proposed algorithms were evaluated in terms of verification. In a verification scenario, a person presents a new signature (probe sequence) and claims to have an identity that exists or not in the system. The validation of the effectiveness of the proposed technique for verification in terms of the rate operating characteristic curves (ROC) is reported in http://www.iti.gr/~djoannid/Gait/index.htm for the USF and the HUMABIO gait databases.

VII. CONCLUSION AND FUTURE WORK

In this paper, a novel gait recognition methodology was presented, based on three new feature descriptors and their fusion using genetic algorithms. Moreover, two new silhouette distributions were introduced that utilize depth information from a stereo camera in order to create 3-D silhouette maps. The proposed system was evaluated with the use of radon transformations and the weighted Krawtchouk moments as feature extractors for identification and verification purposes. By using the best performing proposed methodologies, improvements have been made in recognition performance compared to other methodologies on the reference "Gait Challenge" database. Future work involves the use of new feature descriptors that fully exploit the depth information available.

ACKNOWLEDGMENT

The authors would like to thank Dr. N. Boulgouris for his valuable suggestions at the initial stages of this work.

REFERENCES

[1] M. Murray, "Gait as a total pattern of movement," *Amer. J. Phys. Med.*, vol. 46, no. 1, pp. 290–332, 1967.
[2] M. W. Whittle, "Clinical gait analysis: A review," *Human Movement Sci.*, vol. 15, pp. 369–387, Jun. 1996.
[3] A. Johnson and A. Bobick, "A multi-view method for gait recognition using static body parameters," in *Proc. 3rd Int. Conf. Audio and Video-Based Biometric Person Authentication*, Halmstad, Sweden, Jun. 2001, pp. 301–311.

[4] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, May 2002, pp. 734–742.
[5] L. Wang, H. Ning, W. Hu, and T. Tan, "Gait recognition based on procrustes shape analysis," in *Proc. Int. Conf. Image Processing*, 2002, pp. 433–436.
[6] L. Wang, H. Ning, T. Tan, and W. Hu, "Fusion of static and dynamic body biometrics for gait recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 149–158, Feb. 2004.
[7] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic extraction and description of human gait models for recognition purposes," in *Comput. Vis. Image Understand.*, Apr. 2003, vol. 90, pp. 1–41.
[8] P. J. Phillips, S. Sarkar, I. R. Vega, P. Grother, and K. W. Bowyer, "The gait identification challenge problem: Data sets and baseline algorithm," in *Proc. Int. Conf. Pattern Recognition*, Quebec City, QC, Canada, Aug. 2002, vol. 1, pp. 385–388.
[9] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The human ID gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 162–177, Feb. 2005.
[10] A. Kale, A. Sundaresan, A. N. Rajagopalan, N. P. Cuntoor, A. K. Roy-Chowdhury, V. Kruger, and R. Chellappa, "Identification of humans using gait," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1163–1173, Sep. 2004.
[11] S. D. Mowbray and M. S. Nixon, "Automatic gait recognition via Fourier descriptors of deformable objects," in *Proc. Audio Visual Biometric Person Authentication*, 2003, pp. 566–573.
[12] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "An angular transform of gait sequences for gait assisted recognition," in *Proc. IEEE Int. Conf. Image Processing*, Singapore, Oct. 2004, pp. 857–860.
[13] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "Gait recognition using linear time normalization," *Pattern Recognit.*, vol. 39, pp. 969–979, 2006.
[14] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "Gait recognition using dynamic time warping," in *Proc. IEEE 6th Workshop on Multimedia Signal Processing*, Sep. 29–Oct. 1, 2004, pp. 263–266.
[15] A. Kale, N. Cuntoor, A. N. Rajagopalan, B. Yegnanarayana, and R. Chellappa, "Gait analysis for human identification," presented at the 3rd Int. Conf. Audio and Video Based Person Authentication, Jun. 2003.
[16] Y. Liu, R. Collins, and Y. Tsin, "Gait sequence analysis using frieze patterns," in *Proc. Eur. Conf. Computer Vision*, 2002, pp. 657–671.
[17] P. Salembier, A. Oliveras, and L. Garrido, "Anti-extensive connected operators for image and sequence processing," *IEEE Trans. Image Process.*, vol. 7, no. 4, pp. 555–570, Apr. 1998.
[18] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information," in *Proc. IEEE Intelligent Transportation Systems*, 2001, pp. 334–339.
[19] D. D. Hoffman and M. Singh, "Salience of visual parts," *Cognition*, vol. 63, pp. 29–78, 1997.
[20] K. Moustakas, D. Tzovaras, and M. G. Strintzis, "SQ-Map: Efficient layered collision detection and haptic rendering," *IEEE Trans. Visual. Comput. Graphics*, vol. 13, no. 1, pp. 80–93, Jan./Feb. 2007.
[21] D. Simitopoulos, D. E. Koutsonanos, and M. G. Strintzis, "Robust image watermarking based on generalized Radon transformations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 732–745, Aug. 2003.
[22] P. Daras, D. Zarpalas, D. Tzovaras, and M. G. Strintzis, "Efficient 3-D model search and retrieval using generalized 3-D Radon transforms," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 101–114, Feb. 2006.
[23] N. V. Boulgouris and Z. X. Chi, "Gait recognition using Radon transform and linear discriminant analysis," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 731–740, 2007.
[24] J. Shutler and M. S. Nixon, "Zernike velocity moments for sequence-based description of moving features," *Image Vis. Comput.*, vol. 24, no. 4, pp. 343–356, 2006.
[25] P. T. Yap, R. Paramesran, and S. H. Ong, "Image analysis by Krawtchouk moments," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1367–1377, Nov. 2003.
[26] A. Mademlis, A. Axenopoulos, P. Daras, D. Tzovaras, and M. G. Strintzis, "3D content-based search based on 3D Krawtchouk moments," in *Proc. 3DPVT 2006*. Chapel Hill: Univ. North Carolina, 2006.
[27] J. Little and J. Boyd, "Recognizing people by their gait: The shape of motion," *Videre: J. Comput. Vis. Res.*, vol. 1, no. 2, pp. 1–32, 1998.
[28] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
[29] I. G. Damousis, A. G. Bakirtzis, and P. S. Dokopoulos, "Network-Constrained economic dispatch using real-coded genetic algorithm," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 198–205, Feb. 2003.
[30] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1–3, pp. 7–42, Apr./Jun. 2002.
[31] L. Zongyi and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 6, pp. 863–876, Jun. 2006.
[32] L. Lee, G. Dalley, and K. Tieu, "Learning pedestrian models for silhouette refinement," in *Proc. Int. Conf. Comput. Vis.*, 2003, pp. 663–670.
[33] D. Tolliver and R. Collins, "Gait shape estimation for identification," in *Proc. Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2003, pp. 734–742.

3-D Face Recognition Using Local Appearance-Based Models

Hazým Kemal Ekenel, Hua Gao, and Rainer Stiefelhagen

***Abstract*—In this paper, we present a local appearance-based approach for 3-D face recognition. In the proposed algorithm, we first register the 3-D point clouds to provide a dense correspondence between faces. Afterwards, we analyze two mapping techniques—the closest-point mapping and the ray-casting mapping, to construct depth images from the corresponding well-registered point clouds. The depth images that are obtained are then divided into local regions where the discrete cosine transformation is performed to extract local information. The local features are combined at the feature level for classification. Experimental results on the FRGC version 2.0 face database show that the proposed algorithm performs superior to the well-known face recognition algorithms.**

***Index Terms*—Automatic registration, depth image, local appearance face recognition, 3-D face recognition.**

I. Introduction

Biometric identification is a challenging task that has received a significant amount of interest in the last decades. Among the utilized biometric modalities, the human face is one of the most natural. Moreover, a subject's face images can be acquired easily and unobtrusively. Due to low cost and the wide availability of image acquisition systems, most of the face recognition algorithms are based on 2-D intensity images [23]. However, the algorithms that process intensity images suffer from facial appearance variations that are caused by changes in head pose and illumination conditions. Much effort has been devoted to solving these problems in the 2-D domain. Although significant enhancements have been achieved in the 2-D domain against these variations under controlled conditions, the problem still remains unsolved under uncontrolled, real-world conditions.

Manuscript received May 1, 2007. This work was supported in part by the European Union under the integrated project CHIL, Computers in the Human Interaction Loop under Contract 506909, and in part by the German Research Foundation (DFG) as part of the Collaborative Research Center 588 Humanoid Robots-Learning and Cooperating Multimodal Robots. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Bir Bhanu.

The authors are with the Department of Computer Science, University of Karlsruhe, Karlsruhe 76131, Germany (e-mail: kawagao1979@hotmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2007.902924