INFORMATICA, 2001, Vol. 12, No. 3, 385–412

Feature Matches Filtering Using Geometric Invariants in Image Registration Tasks

Algimantas MALICKAS
Institute of Mathematics and Informatics
Akademijos 4, Vilnius, Lithuania
e-mail: malickas@takas.lt

Rimantas VITKUS
Vilnius University, Čiurlionio 21, Vilnius, Lithuania

Received: January 2001

Abstract. Filtering of feature matches is heuristic method aimed to reduce the number of feasible matches and is widely employed in different image registration algorithms based on local features. In this paper we propose to interpret the filtering process as an optimal classification of the matches into the correct or incorrect match classes. The statistics, according to which the filtering is performed, uses differences of the geometrical invariants obtained from ordered sets of local features (composite features) of proper cardinality. Further, we examine some computationally efficient implementation schemes of the classification. Under the assumption of Gaussian measurement error, the conditional distribution densities of invariants can be approximated by well-known linearization approach. Experimental evidences obtained from fingerprint identification, which confirm viability of the proposed approach, are presented.

Key words: image registration, composite features, geometric invariants.

1. Introduction

The subtask of image registration is common in a variety of applications related to computer vision, such as biometric identification (Ratha *et al.,* 1996), aerial image analysis (Hsieh, 1999), building a mosaic image from a sequence of partial views (Can, 1999), structure from stereo vision (Williamson, 1998), motion analysis (Zheng, 1995), and model-based object recognition (Olson, 1999). Huge amount of publications devoted to this topic indicates that until now there is no single image registration technique suitable for all applications with their different requirements. On the other hand, there is a necessity for unifying the approaches to the image registration, thus creating a regular way to improve a particular registration technique, and help to understand the common points in different image registration algorithms. In this paper we present one of such common approaches, namely, the interpretation of the feature pairs filtering, a supplemental procedure, which could be applied in many different image registration algorithms as statistical classification task.

In our previous paper (Malickas and Vitkus, 1999) we have presented use of the recently proposed feature consensus method (Shekhar *et al.*, 1999) for fingerprint image registration. The method is highly computationally efficient, which is crucial for biometrics applications. We have used classical fingerprint minutia features (Hong *et al.*, 1998) described by position and direction attributes. While in general the transformation relating two fingerprints of the same finger is non-rigid and non-linear, rigid linear transformation was selected as a reasonable starting point. A pair of corresponding oriented points from two images is sufficient to estimate three parameters of rigid transformation, i.e., $2D$ rotation angle and $2D$ translation vector. It appears that performance of the feature consensus with the minutiae features is unacceptable. Only the usage of composite line features, composed of two minutiae features, and application of feature pair filtering with some attributes invariant to application of rigid transformation enables to reach practical acceptability of the registration method in fingerprint identification tasks.

Our search for optimal parameter settings for the fingerprint recognition algorithm, and analysis of approaches to image registration published during the past twenty years, lead us to some generalizations, allowing to distinguish the process of discarding non-corresponding feature pairs using invariants as relatively independent subtask, which may be implemented, or is implemented in some form in the most of existing image registration algorithms. Matches filtering may be formalized as statistical classification task so powerful statistical pattern recognition approaches could be used to optimize it.

2. General Formulation of Feature Pair Filtering Approach

To start, we overview the image registration techniques to which the presented approach could be applied. We are mostly interested in the situations where images are produced by certain imagining devices called cameras. These images represents the $2D$ projections of $3D$ world scenes. The image registration task may arise when there are two images representing two different views of the same scene caused by different location and orientation of the camera relative to the scene. The task is to estimate the geometric transformation that aligns the two images, and depending on the imagining situation, could belong to different groups of $2D$ transformations. Representatives of such groups are the rigid transformation group which consist of $2D$ rotations and translations, the similarity group which is rigid transformation plus uniform change of scale, and the affine transformation group which includes general $2D$ linear transformations plus translations.

Since the image is formed by projecting $3D$ world onto the plain, then transformation groups mentioned earlier only approximately describe the alignment of the two images. Moreover, in some applications the objects represented in the two images may undergo some deformations themselves, or the two images may represent different (but similar in shape) objects. In these more complicated situations the linear transformations could also satisfactorily align the images. While in some medical applications (Lester *et al.*, 1999) it is necessary to estimate non-linear deformation which aligns, for example, two successive slices of the object in $3D$ shape reconstruction.

Another class of applications, which should be considered is the model-based object recognition where the $3D$ model of the object should be aligned to all object instances in the $2D$ image. The alignment transformation belongs to the $3D$-to-$2D$ projection groups, such as orthographic, weak-perspective and perspective projections. Alignment of the $3D$ model to the $2D$ image is a more complicated task, but the techniques used for this task are usually applicable to simpler $2D$-to-$2D$ or $3D$-to-$3D$ alignment tasks.

Above, we have described different situations where the alignment task could arise. Now we shall define the class of alignment algorithms to which the approach presented in this paper is applicable. First of all, the image registration algorithm should be based on local image features. These algorithms employ separate image preprocessing stage to extract the local features, e.g., some distinguished points, edge fragments etc. When the coordinates and other attributes of those features are recorded, the rest of information in the two images is discarded and registration algorithm is based entirely on the two sets of the local image features. The feature extraction stage falls outside the scope of this paper, and we simply suppose that the feature sets of certain type are given in advance. We will call these features as *basic* features. For the sake of simplicity, we will suppose that the image registration algorithm is provided with the homogeneous sets of the features, i. e., all features are described by the same set of attributes and are of the same type, for example, points or lines or straight edge fragments, etc.

The registration algorithms should also use local feature pairs from the two images to estimate the transformation parameters. This implies explicit or implicit recovery of feature correspondence. Thus we should exclude the truly correspondenceless methods, that, for example, use local features, but the transformation parameter recovery is based on the distribution functions of the local feature attributes (Govindu, 1999). Other details of the registration algorithms are mostly irrelevant.

In the following we introduce some definitions. Let $U = \{u_m,\ 1 \leqslant m \leqslant M\}$ be a first set of the basic features. Each feature is characterized by the position measured in a coordinate frame F_1 and possibly by a set of some other attributes. Let $V = \{v_n,\ 1 \leqslant n \leqslant N\}$ be another set of basic features, that are characterized by the position measured in a coordinate frame F_2 and possibly by the same set of attributes as in feature set U. Let $\tau : F_1 \rightarrow F_2$ denote the alignment transformation for two coordinate frames belonging to a transformation group T.

Let us call a pair of the features (u_m, v_n) a *feature match*. A set of feature matches constitutes *matching*. A *correct match* is the match (u_m, v_n) when a feature u_i is correctly matched to its corresponding feature v_j from the other feature set V.

A *composite* feature of cardinality k is an ordered set or k-tuple of k basic features. The composite features are characterized by the attributes of the basic features that constitute the composite feature. But more important is the possibility to compute new attributes, which are *invariant* to the action of the transformation group T. The match $(\boldsymbol{u}_m, \boldsymbol{v}_n)$ of two composite features $\boldsymbol{u}_m = \{u_{m,i},\ i = 1, 2, \ldots, k\}$ and $\boldsymbol{v}_n = \{v_{n,i},\ i = 1, 2, \ldots, k\}$ of cardinality k, is a set $(\boldsymbol{u}_m, \boldsymbol{v}_n) = \{(u_{m,i}, v_{n,i}),\ i = 1, 2, \ldots, k\}$ of the k matches of basic features that constitutes the composite features. It is important to note that the composite feature match is correct if all basic feature matches, which constitute that composite match, are correct.

One more definition is necessary before we proceed further. Estimation of transformation parameters is based on equations relating the attributes of basic features from the different images. Depending on complexity of the basic features and transformation we need some *minimum* number r of basic features to form sufficient number of equations in order to recover transformation parameters. Complexity of the basic features and transformation here is measured by the number of geometric attributes and parameters, respectively. We call composite features having cardinality r as the minimal composite features and the matches of the minimal composite features as the *minimal composite matches*.

Now we can estimate the number of the minimal composite matches, which could be considered in the worst case for recovering the correspondence and estimating the transformation parameters. Let r is a minimum number of basic feature pairs sufficient to estimate transformation parameters. Then from feature set U we can form C_M^r different subsets of size r. The basic features in each subset could be ordered in r factorial $(r!)$ different ways. So the total number of ordered subsets or minimal composite features of cardinality r, that could be formed using basic features from the first set U, is $K = C_M^r r!$. Similarly, $L = C_N^r r!$ is the number of minimal composite features that could be formed from the second basic feature set V. Then, the total number of the minimal composite matches is KL. Let only X $(X < M$ and $X < N)$ basic features has corresponding features in the other image, then $J = C_X^r r!$ composite features of cardinality r are composed exclusively by these X basic features and have corresponding composite feature in the other image. The ratio between the total number of composite feature pairs and the number of pairs representing correct matches is

$$\begin{aligned} R_T &= KL/J = \frac{(X-r)!M!N!}{X!(M-r)!(N-r)!} \\ &= \frac{M\cdot(M-1)\ldots(M-r+1)N(N-1)\cdot\ldots\cdot(N-r+1)}{X(X-1)\cdot\ldots\cdot(X-r+1)}. \end{aligned}$$

This ratio reflects complexity of the correspondence establishment task. Let $\rho = MN/X$ is the ratio between the total number and the number of correct basic feature matches, i.e., ρ represents initial complexity of correspondence establishment between basic features. Then, having in mind, that $r << M$ and $r << N$, $R_T \approx \rho^r$. It means, that correspondence establishment complexity grow exponentially with the growth of composite features cardinality r. Even in the best case, when all features have corresponding ones in other image, i.e., when $X = M = N$ and $\rho = M$, the growth of the ratio R_T is tremendous in typical applications. For example, in our fingerprint identification system the extracted minutiae numbers M or N are scattered in the interval $20 \div 40$, and the value of ρ is distributed in the interval $40 \div 100$. Consequently, in the case of $r = 2$, the ratio R_T can reach ten thousands. Such amount of outliers (false matches) in relation with inliers (correct matches) is too large a burden for the most image registration algorithms. This was proven under the assumption of bounded error in feature attribute measurements for the most popular registration algorithms in model-base object recognition field (Grimson, 1992; Grimson, 1990). And this is why in all proposed systems some techniques of complexity reduction are introduced. Basic feature labeling (Huttenlocher, 1987), grouping (Jacobs, 1989; Lowe, 1987) are the examples of techniques aimed to reduce the total number of hypothetical matches in different recognition systems. While they give promising results, they are rather sophisticated and heuristic (as grouping by Jacobs), or should be implemented in feature extraction algorithms to produce more complex basic features (as various feature labeling techniques).

There is more simple and regular technique for reduction of the number of hypothetical matches based on composite features redundancy. This is the decomposition approach proposed by Olson (Olson, 1999). To explain the approach, suppose that A is a set of all minimal composite matches of cardinality r. The set A may be represented as a sum $A = \cup A_\alpha$ of its subset A_α indexed by all composite matches α of lower cardinality then r, for example $r - 1$. The subset A_α consists of all minimal composite matches that include the composite match α. For example, in the case of cardinality $r - 1$, the subset A_α contains all composite matches $\mu_{ij} = \{\alpha\&(u_i, v_j)\}$, where $i = 1, 2, \ldots, M,\ j = 1, 2, \ldots, N$, and $(u_i, v_j) \notin \alpha$, i.e., each match μ_{ij} contains the same $r - 1$ matches, that constitute composite match α, plus one match from the rest of $KL - r + 1$ possible matches. The correspondence establishment and transformation parameters estimation using only hypothetical matches from subset A_α is considered as solving a sub-problem. This sub-problem is much easier since in A_α there are much less matches then in the whole set A, but the sum of elements in all subsets is larger than the cardinality of the set A, since the subsets A_α partially overlap. The trick is to exploit the redundancy which exists in composite matches. If α is a correct match for some object then this object will be detected when solving the sub-problem A_α. But the same object could be detected when solving another sub-problem A_β. labeled by other correct match β which relates other $r - 1$ features of that object in first image with corresponding features from the second image. Since there is large number of such equivalent sub-problems, it can be used the common randomization technique (Fischler, 1981; Olson, 1997) to limit the number of sub-problems solved while maintaining a low probability of failure.

In the method advocated by this paper reduction of hypothetical matches is achieved by using supplemental filtering process, which is carried out before the main registration algorithm. The matches filtering is based on the feature attributes, called for the sake of simplicity as invariants, that are invariant to the action of the particular transformation group. We will consider only invariants which are functions of the coordinates or other geometrical attributes of basic features. We call them geometrical invariants to distinguish them from the invariants of other modality, e.g., color, texture type, feature type, etc., that could be provided by feature extraction process. In general case the invariants are the functions of the attributes of several basic features. Geometrical invariants for the most interesting transformation groups and types of the basic features are well known (see for example (Weinshall, 1993)). As with the transformation parameter estimation discussed above, for each basic features type there is a minimum number n of basic features necessary to form an invariant function. If we increase this basic feature number n it is possible to form additional invariants, thus by increasing the cardinality of composite features we could obtain more invariant attributes characterizing these features. Moreover, any function of invariants is also invariant to the action of the same transformation group. Consequently there is virtually infinite number of invariants granting possibility to make the best choice for particular tasks.

For example, consider $2D$ space and rigid transformation group, which includes rotations and translations. It is well known that these transformations preserve the distances between the points and the angles between the lines. Thus, if we have composite feature constituted by two points we can calculate one invariant distance between the points, if we have three points then we can calculate three distances using three pairs of points, and three angles using three pairs of lines connecting the points. Moreover, we can compute the area of the triangle, which is also invariant to the action of rigid group, but we need all three points to calculate that invariant. If we have the so-called oriented feature points, i.e., points additionally characterized by direction angle, then having composite feature of cardinality 2 we could calculate not one, but three useful invariants: the distance between the points and the two angles between the line connecting points and the direction vectors of the two points.

Thus, if we have the correct match of two corresponding composite features, then the same invariants calculated from the two features should coincide, provided the measurements are done without errors. In case of the measurements with errors, the calculated invariants from the corresponding features would be equal with some degree of uncertainty. Having a proper model of propagation of uncertainty from the attribute measurement error to invariant calculations, one can introduce invariant similarity measure, and use it to discard the incorrect matches in order to reduce the ratio of false to correct matches.

Below we describe a regular way to apply this scheme. Let the cardinality of minimal matching is r and there is no invariants to reduce the outlier/inlier ratio, or the invariants calculated from the minimum composite features are insufficient to reduce this ratio to an acceptable degree. Then it is possible to consider composite features and matches having $r + 1$ or, more generally, $r + c$ cardinality. The composite features of higher cardinality inherits the invariants from features of lower cardinality, yet they give opportunity to compute other invariants, which can serve as additional constrains to detect and discard more incorrect matches.

In fact the filtering could be regarded as a traditional statistical classification task. Suppose we judge about the similarity of two invariant values looking if their difference is close to zero. The difference (or differences of several invariants) is random variable(s). The conditional distribution of invariant difference in case of correct match and false match are quite different as would be demonstrated in the following sections. Thus using the differences of appropriate invariants obtained from a composite feature match we can assign this match to correct or false match class in some statistically optimal way.

3. Two Ways of Incorporation of Matches Filtering

There are essentially two practical ways to incorporate the feature match filtering into the registration algorithms. Both have their implementation drawbacks and advantages.

The first way is to use the higher cardinality composite feature matches exclusively for filtering purposes in order to discard the incorrect matches of lower cardinality. This is based on the fact that the composite feature match of cardinality k can be correct only if all k basic feature matches, which constitute the composite match are correct. Thus, if we take any subset of lower cardinality $l < k$, then this composite match also would be correct, if the composite match of cardinality k is correct. In this form the matches filtering could be adopted by any registration algorithm with minor modifications of the algorithm.

We shall describe the filtering process for the case when composite matches of cardinality $r + 1$ are used to filter the composite matches of cardinality r. Generalizations are straightforward. Suppose we have K and L composite features of cardinality r in the two feature set. The total number of hypothetical matches is equal to KL. Let the construction of composite features of cardinality $r + 1$ is done by adding to the end of each composite feature of cardinality r an additional basic feature. Thus from each composite feature of cardinality r we could construct $M - r$ and $N - r$ new composite features of cardinality $r + 1$ in first and second sets respectively. The total number of hypothetical matches of cardinality $r+1$ will be approximately equal to $KLMN$, where M and N are the numbers of basic features in two images. Then the filtering process could be produced as follows. Let $\{\alpha_p\}$ is a set of invariants, that could be calculated using composite feature of cardinality $r+1$. And suppose we have the procedure according to which the composite matches of cardinality $r + 1$ are assigned into correct and incorrect match class. The number and exact form of the invariants and specific classification procedure at this moment are unimportant. Let each cell of $2D$ accumulator array $A = \{a_{k,l},\ k = 1, 2, \ldots, K;\ l = 1, 2, \ldots, L\}$ is associated with one of KL composite matches of cardinality r. The array is initialized by zeroes. Recall, that each composite feature of cardinality $r+1$ is an extension of some composite feature of cardinality r and therefore each composite match of cardinality $r+1$ is related with the match of cardinality r from which it was constructed. As was discussed earlier, if the match of cardinality $r+1$ is assigned to correct match class, then related match of cardinality r should also be considered as correct match. So each correct match of cardinality $r + 1$ casts a vote for its related match of cardinality r into the associated accumulator element. When voting procedure is finished, the correct matches of cardinality r are the matches that are associated with non zero accumulator elements. And only those matches should be used in the subsequent registration procedure. It should be mentioned that in most existing recognition systems based on geometric hashing this filtering is inefficient (Grimson, 1991). Filtering appears to be more effective if we decide that match of cardinality r is correct when associated accumulator element exceeds non zero threshold. But the selection of this threshold is out of scope of this paper.

The most known algorithms which use such approach are object recognition algorithms based on indexing or geometric hashing scheme. (Hummel, 1988a; Lambdan, 1988a; Lambdan, 1988b; Costa, 1990; Tsai, 1993; Tsai 621). In fact, these algorithms use the feature matches filtering together with decomposition into the sub-problems and very efficient way of implementation, which will be described below. Computational efficiency mades this scheme very popular in the object recognition field.

We illustrate the scheme for $2D$ basic feature points and affine transformation group. In order to recover the $2D$ affine transformation, it is sufficient to have three corresponding point matches. Using three points we can form the affine coordinate system and represent all other feature points by their two coordinates in this frame. For example, if we take three non collinear points x_1, x_2, x_3 in the plain and place zero point of coordinate system in x_1, then any point x of the plain can be represented uniquely by two coordinates (a, b):

$$x = a(x_2 - x_1) + b(x_3 - x_1).$$

It is easy to show, that if we apply the affine transformation τ to all points $x' = \tau x$, and again select the same three transformed points x'_1, x'_2, x'_3 to form the basis in transformed plain, then affine coordinates of any other transformed point x' do not change:

$$x' = a(x'_2 - x'_1) + b(x'_3 - x'_1).$$

Three basis points plus a fourth point constitutes a composite feature of cardinality 4. The affine coordinates of fourth point are two invariants, that are used to filter hypothetical composite matches of cardinality 4. But to recover the $2D$ affine transformation parameters only three pairs of corresponding points are neccessary. So in geometric hashing schemes the composite matches of cardinality 4 are used to filter matches of triplets forming the basis. Note, that efficient implementation scheme described in following section introduces some discretization errors.

In the following we demonstrate how to incorporate the hypothetical matches filtering into the registration algorithm (Gold, 1998) from rather different class of algorithms that uses some global optimization techniques. We have implemented this algorithm anticipating that it will facilitate the analysis of our fingerprint image registration algorithm, since it is too slow to be used in biometrics applications. This is an elegant algorithm, which uses deterministic annealing to escape from local minimum, and softassign technique to satisfy the assignment matrix constraints, and simultaneously recovers feature correspondences and transformation parameters. The algorithm minimizes the following objective function:

$$E(m, T, R) = \sum_{j=1}^{M} \sum_{k=1}^{N} m_{jk} \left\| u_j - T - Rv_k \right\|^2 - \alpha \sum_{j=1}^{M} \sum_{k=1}^{N} m_{jk},$$

where T and R denotes $2D$ translation and rotation transforms, respectively, and coefficients m_{jk} represents feature correspondences ($m_{jk} = 1$, if the feature point u_j corresponds the feature point v_k, otherwise $m_{jk} = 0$). Although some examples presented in the literature (Gold, 1998) look similar to the fingerprint minutiae point matching, nevertheless, we have not succeeded in selecting the proper parameter values, and the algorithm always converges to spurious local minimum, except for the cases when the two minutiae sets coincide. The reduction of hypothetical matches should influence favorably the algorithm, or at least its convergence rate, and speed it up by excluding large amount of calculations. The feature match filtering could be realized in very similar manner to that described in the beginning of this section. Let the matrix $\{n_{jk}\}$ having the same size $M \times N$ accumulates the votes for minutiae features matches. Using the three invariants mentioned above (Malickas and Vitkus, 1999), we can classify the composite matches of cardinality two into the correct and incorrect ones. The composite match of cardinality 2 is an ordered set of only two minutiae matches. So the correct composite match votes for one of those two minutia matches (let it be the first minutiae match) as feasible match, and adds the votes into the appropriate element n_{jk}. When all votes are casted, the elements n_{jk} that exceed some threshold will be equaled to 1, otherwise to 0. Formally, to incorporate the filtering into the optimization algorithm, we should replace the term m_{jk} in the objective function by the product $m_{jk}n_{jk}$. In practice the algorithm should simply check the element n_{jk} before it produces any calculations related to the m_{jk} term. The algorithm should skip all calculations and assign zero to the term m_{jk}, if $n_{jk} = 0$.

The other way to incorporate the filtering is to modify the registration algorithm for using the composite feature of higher cardinality not only in filtering stage, but also in a further registration procedure. It should be noted that such modification is always possible, since composite features of lower cardinality, which were used in the original registration algorithm, bear less information, yet are sufficient to recover the transformation parameters. Thus, composite features of higher cardinality, which include all information of the lower cardinality features, will provide even more information to the modified algorithm.

The main disadvantage of this approach is the exponential growth of hypothetical matches which will affect not only the filtering stage, as in the previous approach, but also the main registration algorithm which may be more computationally demanding. Suppose, we use composite features of cardinality $r + 1$, instead of using composite features of cardinality r. Then the number of composite features will increase by the factor $M - r$ and $N - r$ in the first and second image, respectively, and the total number of the hypothetical matches will approximately increase by the factor MN. Consequently, using this approach it is an urgent necessity to incorporate the efficient implementation algorithms described in the next section. For the algorithms which use computationally expensive global optimization, similar to those described in the previous paragraph, this approach could be unacceptably slow.

One of the advantages of this approach is the possibility to employ the additional information of composite features of higher cardinality. For example, if we have more than minimum number of points or other features, we can recover the transformation parameters more precisely.

To illustrate the second approach, we present the modification of the registration algorithm based on Hough transform (Ratha *et al.*, 1996), which we have implemented in order to compare its efficiency with feature consensus algorithm (Malickas and Vitkus, 1999). The original Hough algorithm have been implemented using the oriented point features. We use three-dimensional accumulator array with one dimension for rotation and two for translation parameters. For each minutiae pair and each rotation angle, which aligns the minutia directions within some bounded error, we calculate translation vector which aligns these minutia. Then we increment the accumulator array element pointed to by the angle and translation vector as well as nine neighboring elements in the translation plane of the accumulator. The implementation using composite features of cardinality two differs only in two aspects: the minutia direction is replaced by the direction of a line fragment connecting the two minutia, and the minutiae location is replaced by the location of the fragment center. Nevertheless, as mentioned earlier, in the case of composite features we have the possibility to use additional information. For example, the error of the fragment angle depends on the distance between the minutia, in particular error grow when distance gets smaller. The angles between the corresponding minutiae directions could also be used as rotation angle estimates and these estimates have stable error. So it is possible to combine these three estimates (i.e., take a weighted average) in order to reduce the estimation error.

4. Efficient Implementation Algorithms

It is necessary to find answers to some questions which arise after closer examination of the approach, considered above. First of all, the relation between the outlier/inlier ratio ρ for basic feature matches and R for composite feature matches of cardinality r is expressed as $R \approx R_T \approx \rho^r$. Consequently R grow exponentially with r and it is unclear do the matches filtering could eventually reduces ratio R below the ρ. We do not have formal proof of this assertion in general, but it will be addressed in the following sections of this paper from the statistical point of view. Anyway, the answer is influenced by the two circumstances. First, the number of possible invariants increases very rapidly with r, thus the potential number of constraints on the correct matches also increases with r. Second, the tightness of these constraints heavily depends on the degree of uncertainty involved in the invariant value calculation.

The next problem is rapid grow of computational load with the increase of r, and how this combinatorial explosion should be managed. Answer to this question, or at least partial answer will be given in this section.

First of all, the large redundancy in composite features should be emphasized. For example, each basic feature from the first set U has at most only one corresponding basic feature in second set V, but the same corresponding pair of basic features emerges in a large number of correct composite matches. To be more precise, let X be the number of correct basic feature matches, and r is the cardinality of the composite features. Then each correct basic feature match could be found in $Y = C_X^{r-1} r!$ correct composite matches. This can be enormous number, while in order to estimate transform parameters it is enough to catch only one correct match of cardinality r_1, where $r_1 < r$. Most methods discussed below exploit this redundancy.

The simplest way to limit the number of considered composite features is to set some regular constraints on their formation. For example, we can form the pairs of minutiae only if the distance between them is within certain bounds. This is justified not only by the reduction of total number of matches. It is well known, that some configurations of basic features can produce large errors in invariant value calculations even when moderate errors are produced in the basic feature attribute measurements (Costa; 1990). Usually, there exist numerical values indicating these unstable configurations, which could be used to prevent the formation of unwanted composite features (Costa, 1990; Olson, 1992).

The randomization technique is other regular method widely used to limit the number of hypothetical matches, as well as in other applications for the same aim (Bergen, 1991; Xu, 1990). According to it only a limited number K of randomly selected matches are considered (Fischler, 1981; Olson, 1997). It is possible to select the number K in a such manner, that the probability of failure is reasonably small.

The next method described bellow is according to our knowledge new, and quite general one for dealing with the combinatorial explosion. It was not implemented practically, thus only the theoretical framework will be presented. It is likely that this method could be applied together with the previously outlined methods.

As mentioned earlier, the composite features are ordered sets or k-tuples of basic features. If we have a set of the k basic features from the first image, we can form $k!$ ordered sets or composite features from it. Thus, the number of ordered feature sets of the size k is $k!$ times larger then the number of unordered feature sets. Suppose, that all k basic features have there corresponding basic features in the second image. Then for each of the $k!$ ordered sets of basic features from the first image there exist a corresponding ordered set in the second image. Consequently, from one pair of corresponding unordered sets we get $k!$ pairs of corresponding ordered sets or correct matches. But the registration algorithm do not know in advance the correspondence of the basic features and should consider possible matches of all composite features from the first image with all composite features from the second image. The total number of possible matches of the composite features formed from the pair of corresponding unordered sets is $k! \times k!$. So in addition to $k!$ correct matches we should consider $k! \times (k! - 1)$ incorrect matches. The idea of the method is to recover correspondence of the unordered sets of basic features instead of ordered ones.

To establish the correspondence between unordered sets of k features we should have specific invariant α, which is the function of all k features (more precisely α is a function of some attributes of all k features) and its value does not depend on the order of its arguments, i.e., α is symmetric function of its arguments. Such invariants do exist at least for some widely spread situations. For example, let consider rigid transformation group on the plain, which preserves the area of geometric figures. Let $k = 3$ and the basic features are $2D$ points. Then the area of a triangle build by three points is an invariant, which does not depend on the sequence of its apexes, and thus is symmetric function of the triangle apex points.

If by comparison of appropriate invariants the pair of unordered sets was judged as incorrect match, then incorrect will be all $k! \times k!$ matches of ordered sets. If the pair was not discarded as incorrect, then the correspondences between the basic features inside the unordered sets should be established. For this purpose a system of partially ordered invariants may be used. The partial order is based on the number of features which was used to calculate the invariant, i.e., for two invariants the relation $\alpha_1 > \alpha_2$ is true if α_1 is a symmetric function of the set D_1 of feature attributes, and α_2 is a symmetric function of the set D_2 of feature attributes, where D_2 is a subset of D_1. If comparison of the first pair of invariants does not discard unordered match, the next invariant according to the partial order should be considered. An interpretation of the next step is the following. We withdraw one feature from the first and one feature from the second feature set. And exactly in the same way as with the original feature sets, the second level invariant tests the compatibility of two unordered sets of features of cardinality $k-1$. Thus, in the next level the $(k-1)! \times (k-1)!$ matches of ordered sets of cardinality k can be rejected. This hierarchical process can obviously be repeated until the correspondence will be established between all k basic features in the sets or in some step all possible ordered matches will be discarded.

The last technique called indexing or geometric hashing is very popular in model-based object recognition, but is also well suited for biometrics applications. Since the object models are accessible before the recognition process, large number of calculations on the data can be performed in advance. In the preprocessing stage the models of several objects are processed in sequential order. For each model a different feasible composite features of cardinality r are selected as a basis. Then, the invariant representation of all other features of the model are calculated in terms of the current basis. In example presented in previous section the invariant representation of the points were two affine coordinates. These invariants or their functions are used as indices to hash table and the information about model and the current basis is stored into the pointed location of the table. If several sets of indices point to the same bucket of the hash table, the information about all objects and their appropriate basis are stored as a list.

In the recognition stage features are extracted from the image, and the composite features of cardinality r are selected as the basis in sequential manner. For each other basic feature of the image the same invariant representation according to the current basis is calculated and used to index the hash table. All the model-basis pairs are retrieved from the indexed bucket and their individual scores are incremented. When the invariant representations of all features with respect of current basis are used, the model-basis pair with the largest score is found. If largest score exceeds a predefined threshold, the existence of the model instance in the image is hypothesized. Then the process is repeated with the next basis.

The feature matches filtering is produced at the entry retrieval from the hash table. This is a highly efficient technique, since the selected basis is matched against all models and their basis simultaneously. On the other hand quantization of the invariants space introduces some errors into the filtering process.

5. Estimation of the Probability Densities of the Invariants

In the next two sections we formulate feature matches filtering as an optimal statistical classification procedure. Differences between invariants will be the statistics according to which feature matches are classified into the correct and incorrect ones.

Learning is an important step in the statistical classification procedure, during which the conditional probability densities of statistics are estimated. In our case this means recovery of conditional probability densities of invariants differences for correct and false match classes. Once these densities are estimated, the traditional statistical decision theory could be applied in order to classify the samples. If, for some training set of image pairs, correspondences between features are established manually, then it is possible to use some standard learning algorithms to estimate these conditional probabilities. More simple way is to approximate the conditional distributions using measurement error distributions, since invariants are the functions of measured feature attributes. Usually, the manual establishment of correspondences between the features is also necessary in order to estimate measurement error distribution. Nevertheless estimation of measurement error is easier task and if we estimated this distribution, we could apply it to approximate the distributions of variety of invariants. This is the way we propose to use for feature matches filtering.

The measurement errors of the feature attributes is usually approximately Gaussian and independent. For example, in the case of fingerprint minutiae, the position and orientation angle measurement errors are close to Gaussian as will be shown empirically in Section 7. In this case, it is possible to approximate the errors of the invariants by Gaussian distribution. This is done by the standard linearization technique (Tsai 621; Costa, 1990). Suppose the invariant $\alpha = f(X)$ is a function of the statistically independent feature attributes vector $X = \{x_1, x_2, \ldots, x_n\}$ and the attribute x_i is distributed according to Gaussian distribution with parameters $(\bar{x}_i, \sigma_i)$. Then the invariant α is also random variable and its distribution could be approximated by Gaussian distribution with mean and variance $(\bar{\alpha}, \sigma_\alpha)$ (Costa, 1990):

$$\bar{\alpha} = f(\bar{X}), \tag{1}$$

and

$$\sigma_\alpha^2 = \sum_{i=1}^{n} \left(\frac{\partial}{\partial x_i} f(\bar{X}) \right)^2 \sigma_i^2. \tag{2}$$

If we have two invariants $\alpha_i = f_i(X)$, $i = 1, 2$, then the covariance is equal to

$$\sigma_{\alpha_1 \alpha_2}^2 = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_1(\bar{X}) \frac{\partial}{\partial x_i} f_2(\bar{X}) \sigma_i^2. \tag{3}$$

Since multidimensional Gaussian distribution is fully defined by the moments up to second order, having these parameters it is easy to construct the joint Gaussian distribution of any set of invariants.

Below we present the expressions of means and variances for invariants used to discard incorrect matches in our image registration algorithm. As was mentioned above the basic features were minutia or oriented points, that are characterized by three attributes:

their coordinates (x, y) and direction φ_m. In algorithm we use composite features of cardinality 2, i.e., a pair of minutia characterized by their coordinates (x_1, y_1), (x_2, y_2) and directions φ_{m1}, φ_{m2}. In addition to these attributes each minutia pair could be characterized by three attributes that are invariant to action of the rigid transformation group. These are the invariants used in filtering process: distance between minutiae d and two differrences of angles $\Delta\varphi_1 = \varphi_l - \varphi_{m1},\ \Delta\varphi_2 = \varphi_l - \varphi_{m2}$, where φ_l is a slope angle of line connecting two minutia. The expression for the distance d and its partial derivatives are:

$$\begin{aligned} &d = f_1(X) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \\ &\frac{\partial}{\partial x_i} f_1(\bar{X}) = (-1)^{i-1}(\bar{x}_1 - \bar{x}_2)/\bar{d}, \\ &\frac{\partial}{\partial y_i} f_1(\bar{X}) = (-1)^{i-1}(\bar{y}_1 - \bar{y}_2)/\bar{d}, \end{aligned} \tag{4}$$

where $i = 1, 2$. Let σ_x, σ_y be the standard deviations of x and y coordinates of the minutiae. Then by substituting (4) into the (2) we obtain the variance σ_d^2 of distance d:

$$\sigma_d^2 = \frac{2}{\bar{d}^2}\left((\bar{x}_1 - \bar{x}_2)^2\sigma_x^2 + (\bar{y}_1 - y_2)^2\sigma_y^2\right).$$

If $\sigma_x \approx \sigma_y = \sigma_c$, then $\sigma_d^2 = 2\sigma_c^2$ and σ_d^2 does not depend on the size of d.

The distributions of other two invariants are identical, so we will derive distribution parameters only for first of them:

$$\Delta\varphi_1 = \varphi_l - \varphi_{m1},$$

where φ_{m1} is Gaussian variable obtained from the feature extraction process with average value $\bar{\varphi}_{m1}$ and variance σ_{m1}^2. The φ_l is calculated as

$$\varphi_l = f_2(X) = \operatorname{arctg}\left(\frac{y_1 - y_2}{x_1 - x_2}\right).$$

The partial derivatives of φ_l are:

$$\begin{aligned} &\frac{\partial}{\partial x_i} f_2(\bar{X}) = (-1)^i \frac{1}{1 + \left(\frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1 - \bar{x}_2}\right)^2} \frac{\bar{y}_1 - \bar{y}_2}{(\bar{x}_1 - \bar{x}_2)^2} = (-1)^i(\bar{y}_1 - \bar{y}_2)/\bar{d}^2, \\ &\frac{\partial}{\partial y_i} f_2(\bar{X}) = (-1)^{i-1} \frac{1}{1 + \left(\frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1 - \bar{x}_2}\right)^2} \frac{1}{(\bar{x}_1 - \bar{x}_2)} = (-1)^{i-1}(\bar{x}_1 - \bar{x}_2)/\bar{d}^2. \end{aligned}$$

Then from (2) we obtain the variance σ_φ^2 of φ_l:

$$\sigma_\varphi^2 = \frac{2}{\bar{d}^4}\left((\bar{x}_1 - \bar{x}_2)^2\sigma_x^2 + (\bar{y}_1 - \bar{y}_2)^2\sigma_y^2\right).$$

If $\sigma_x \approx \sigma_y = \sigma_c$, then $\sigma_\varphi^2 = 2\sigma_c^2/\bar{d}^2$ and depends on the size of distance d between te two minutia.

Finally, the invariant $\Delta\varphi_1$ is a sum of two independent Gaussian variables so it is Gaussian random variable with average value $\bar{\varphi}_l - \bar{\varphi}_{m1}$ and variance $\sigma_l^2 + \sigma_{m1}^2$.

Since d and φ_{m1} are independent, the covariance between d and $\Delta\varphi_1$ is equal to covariance between d and d and φ_l. From equation (3) this covariance is equal to

$$\sigma_{d\varphi}^2 = 2\frac{\bar{y}_1 - \bar{y}_2)(\bar{x}_1 - \bar{x}_2)}{\bar{d}^3}(\sigma_y^2 - \sigma_x^2).$$

It can be seen, that covariance is equal to zero if the variances of x and y coordinates are equal.

6. Feature Matches Filtering as Optimal Classification

In the previous section we demonstrated the possibility to approximate the distribution densities of invariants by a Gaussian distribution using the first order approximation under the assumption that measurement errors are independent and Gaussian. Having these distributions of the invariants we can propose filtering scheme based on statistically optimal classification procedure. We start from classification using one invariant.

The task can be formulated as follows. Having a pair of composite feature (u_i, v_j) from the first feature and the second feature sets, and respective invariants α_i and α_j, we should decide if this pair is a correct or a false match. The difference $\Delta\alpha = \alpha_i - \alpha_j$ will be used as a statistic in the decision process. When the pair (u_i, v_j) is a correct match, invariants α_i and α_j are independent and identically distributed Gaussian random variables with the distribution function $N(\bar{\alpha}, \sigma_\alpha)$ as shown in previous section. Then $\Delta\alpha$ is distributed according to Gaussian distribution $N(0, \sqrt{2}\sigma_\alpha)$.

When the pair (u_i, v_j) is an incorrect match, we suppose that α_i and α_j are independent and identically distributed random variables having uniform distribution function. This assumption seems to be reasonable, for example, in object recognition applications (Costa, 1990) and the experimental evidences for this model in the fingerprint identification field will be given in the next section. Also other distribution models, like truncated Gaussian distribution could be used, if necessary, with minor changes in the procedure. Let W be the width of the interval in which α_i and α_j are distributed. Then in the case of false matches $\Delta\alpha$ is distributed according to the triangle distribution, which is non zero in a finite interval:

$$t(x, W) = (1 - |x|/W)/W, \quad \text{for} \quad -W < x < W.$$

This follows from the well known fact that the distribution density f_{diff} of difference of two random variables distributed according to densities f_1 and f_2 is a convolution $f_{diff}(x) = f_1(x) * f_2(-x)$. This is true for invariants like distances or areas. For angle invariants distributed according to uniform distribution the distribution of the angle difference is also uniform since convolution is cyclic.

From the viewpoint of statistical inference, we should select between two simple hypothesis: is $\Delta\alpha$ distributed according to $N(0, \sqrt{2}\sigma_\alpha)$ (the match is correct – null hypothesis H_0), or according to the triangle distribution $t(W)$ (the match is incorrect – alternative hypothesis H_1). The disappointing circumstance here is that both distributions are unimodal, symmetric, and centered at the same zero point. Nevertheless, the hypotheses can be well discriminated at the tails of distributions provided that W is much larger than $\sqrt{2}\sigma_\alpha$. In order to simplify the situation we standardize the invariants dividing them by standard deviation $\sqrt{2}\sigma_\alpha$. In the correct match case this operation changes distribution to $N(0, 1)$. In the false match case the parameter W of triangle distribution function also should be replaced by $W' = W/\sqrt{2}\sigma_\alpha$, i.e., the width is measured in standard deviation units.

The standard decision rule (Levin, 1989; Fomin, 1986) is to form likelihood ratio $l(\Delta\alpha) = f(\Delta\alpha|H_1)/f(\Delta\alpha|H_0)$ and compare it with the threshold c. If $l(\Delta\alpha) < c$ then the null hypothesis (correct match hypothesis) is accepted, otherwise the alternative hypothesis is accepted. The threshold value depends on the accepted decision criterion. Having in mind the substitutions made in previous paragraph, the likelihood functions $f(\Delta\alpha|H_0)$ and $f(\Delta\alpha|H_1)$, which are simply conditional densities of random variable $\Delta\alpha$, can be expressed as:

$$f(\Delta\alpha|H_0) = k_0 \mathrm{e}^{-(\Delta\alpha)^2/2},$$
$$f(\Delta\alpha|H_1) = k_1(1 - |\Delta\alpha|/W'),$$

where k_0 and k_1 are normalizing constants. The same decision will be made if we compare $\log(l(\Delta\alpha))$ with $\log(c)$.

Some aspects of this decision rule should be discussed. Since both likelihood functions are symmetric, only positive argument can be considered. First, it should be noted, that $l(\Delta\alpha)$ is not monotonic and topology of likelihood curve intersections depends on the W' size. For example, for W' values that are larger than some value between 2 and 3, the straight line of triangle density intersects the standard Gaussian density at two points. Nevertheless it is obvious, that the critical zone for rejection of null hypothesis should extend from the point defined by threshold to infinity.

More difficulties causes selection of classification criterion which defines the threshold c. It seems, that the Neuman–Pearson or maximum likelihood ratio criterion is acceptable. The more general Bayesian decision rule could be applied also, but we face the problem of selection of right penalties for wrong decision. For example, according to maximum a posteriori probability criterion the penalties are equal in both cases: 1) if we assign the match to correct match class when in fact it is false match; and 2) if we assign the match to false match class when in fact it is correct match. The threshold then is equal to the ratio of prior probabilities of classes. As it is easy to estimate from the correct and incorrect match numbers provided in Sections 2 and 3, usually this ratio is less then 1/50. On the other hand for the typical values of W' which spread in the range 5–30, the likelihood ratio is always larger the 1/50. Consequently according to this decision rule we always should assign the matches to the false match class. Nevertheless, from the practical point of view selection of the criterion is not a big problem. If we have some measure of quality for registration or for the whole identification or object recognition process, then we could experimentally find maximum value of this measure depending on the threshold value used in filtering algorithm.

In the case of several invariants, the decision making process remains unchanged except for the computation increase. Using the method outlined in the previous section, it is possible to estimate the covariance matrix Σ, and the likelihood function in case of correct matches becomes as

$$f(\Delta\alpha|H_0) = k_0 \exp\left(-\boldsymbol{\Delta\alpha}^T \Sigma^{-1} \boldsymbol{\Delta\alpha}/2\right),$$

where $\boldsymbol{\Delta\alpha}$ now is a vector. The likelihood function in the case of alternative hypothesis H_1 is a product of one dimensional density functions:

$$f(\boldsymbol{\Delta\alpha}|H_1) = k_1 \prod_i (1 - |\Delta\alpha_i|/W_i'),$$

since in this case the differences of the invariants are independent. While it is possible to use efficient look-up table methods to compute $f(\boldsymbol{\Delta\alpha}|H_1)$, nevertheless, two different types of densities in likelihood ratio complicates its calculations. As it will be shown experimentally in the next section, the distribution function of invariants could be approximated by truncated Gaussian distribution. In this case the logarithm of likelihood ratio could be expressed as follows:

$$\ln\left(F(\boldsymbol{\Delta\alpha}|H_1)/F(\boldsymbol{\Delta\alpha}|H_0)\right) = (\boldsymbol{\Delta\alpha}^T)\mathbf{A}(\boldsymbol{\Delta\alpha}) + C,$$

where $\mathbf{A}$ is the matrix, pre-computed from the covariance matrices, and C is a constant. Thus, in order to make decision, it is necessary to calculate quadratic form and compare its value with some threshold.

7. Experimental Data

In this section we will demonstrate experimentally the validity of general assumptions about the prior distributions of the invariants and the effectiveness of the feature matches filtering framework in the fingerprint identification task. We start with the estimate of the feature attribute measurement error distributions, and prior distributions of the invariants. These error may be large, since it is clear in advance that the rigid transformation group do not describe exactly the transformation between the two fingerprint images. In general case, only nonlinear elastic deformations could align exactly the two fingerprints.

7.1. *Prior and Conditional Distributions*

As mentioned earlier, in our fingerprint identification system we use minutiae features, characterized by their location and direction attributes. In order to evaluate the prior and conditional distributions of these attributes, it is necessary to establish the true correspondence between the minutiae in different fingerprints. To eliminate the feature extraction process, which is not under consideration in this paper, the true correspondence was established not using all minutiae in two fingerprints, but only the minutiae features that have been automatically extracted and presented for identification process. The correspondence have been established between each pair of five fingerprints from five fingers, totally 50 different fingerprint pairs having 853 correct pairs of minutia.

The correspondences between minutia have been settled by neutral person not skilled in fingerprint identification. The pairs of preprocessed binary fingerprints were presented on a monitor screen. The extracted minutiae features were marked and enumerated, so that the corresponding minutia indices could be recorded.

Since the minutia matching is tedious and long process, some system to check matched pair consistency have been developed. First of all, each pair of fingerprints had been submitted for matching twice with exchanged positions. The operator, which conducted the matching was asked to find for each minutia from the left side fingerprint the corresponding minutia in the right side fingerprint if it exist. Then the two lists of matched pairs were compared to check if they contain the same matched pairs. If any discrepancy was found, the operator was asked to reestablish the appropriate correspondence.

The next step was to check the consistency of the matched pairs between more than two images. Suppose we have N fingerprint images of the same finger and from each fingerprint image I_n the feature extraction process extracts L_n minutiae features, $n = 1, 2, \ldots, N$. Let the fingerprint images and minutia inside the fingerprint images are somehow indexed. Then we can form the set $K = \{m_{n,k},\ n = 1, \ldots, N,\ k = 1, \ldots, L_n\}$ of all minutia from all fingerprints, that are indexed by its fingerprint image index and its order index inside the fingerprint. The correct correspondence process should to each finger minutiae assign a subset of K which consist at most of N corresponding minutia, one from each image. The subsets corresponding to different minutia must not overlap since each image minutiae should match only one minutiae on the finger. In fact, the correspondence induces the equivalence relationship between the elements of the set K and divides them into non overlapping equivalence classes.

The consistency checking was based on this equivalence relationship and the equivalence classes formation process. Namely, if the established correspondences are correct, then starting from any single minutiae from the particular equivalence class we should form exactly the same class. The process of the equivalence class formation is as follows. For each pair of the fingerprint images we have a list in which the corresponding minutia indexes are recorded. The minutia from each fingerprint are recorded in $N - 1$ lists. Let us start from the minutiae $m_{n,k}$ and form the equivalence class $A_{n,k}$ by scanning appropriate $N - 1$ lists of matched minutiae pairs and adding to equivalence class $A_{n,k}$ all minutiae that were paired with $m_{n,k}$. Then we form the equivalence classes using all other minutia from $A_{n,k}$ as the starting point. The correspondences related with $A_{n,k}$ are established correctly if all new classes coincide with equivalence class $A_{n,k}$. The rigid transformation parameters for each pair of fingerprint were estimated by the least square method using corresponding minutia. After applying the transformation to

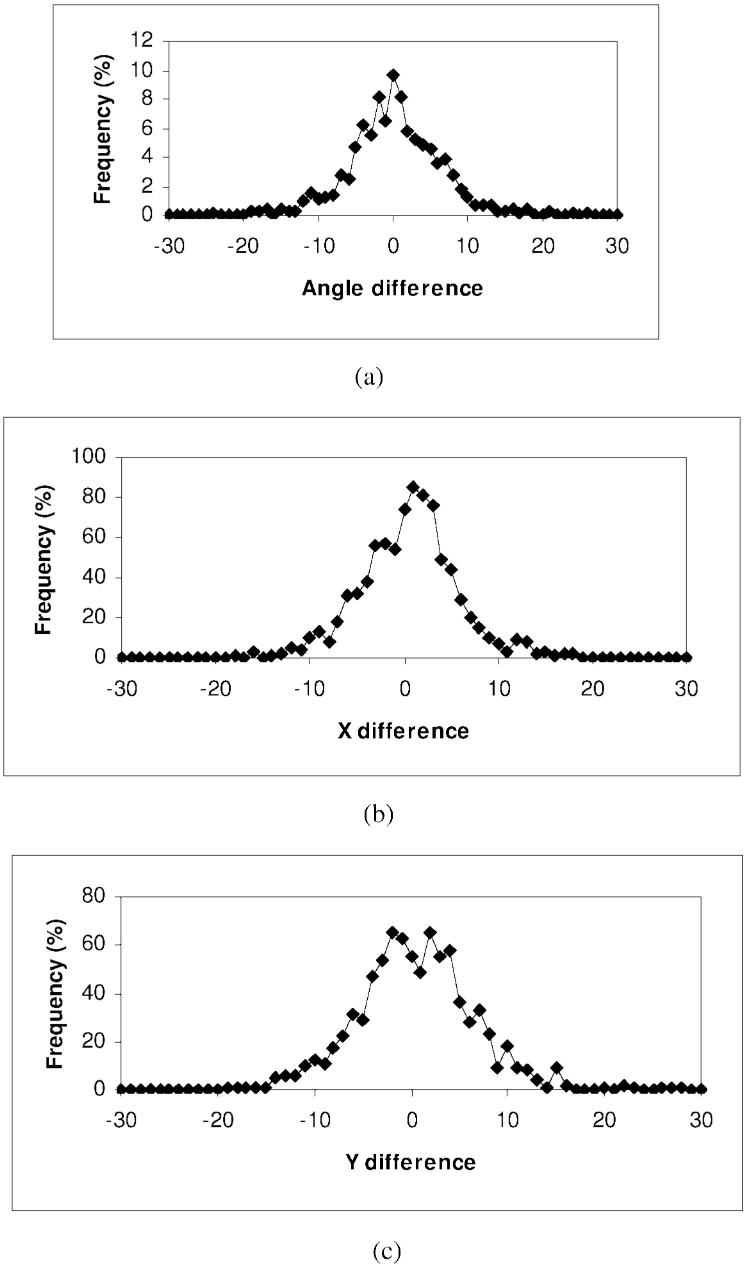
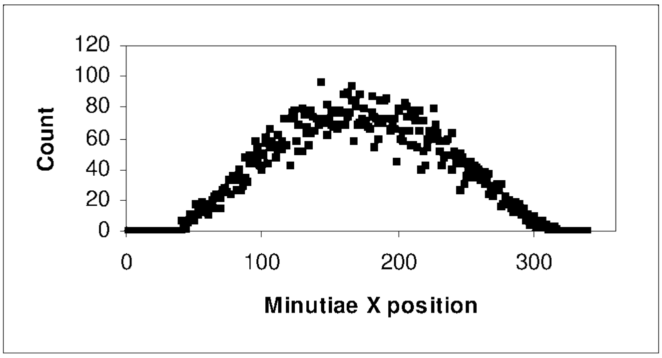
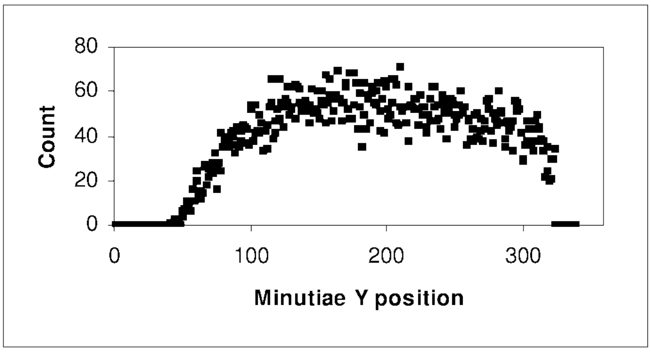


Fig. 1. The distributions of attributes measurement errors: (a) minutia orientation angles, (b), (c) x and y coordinates respectively.

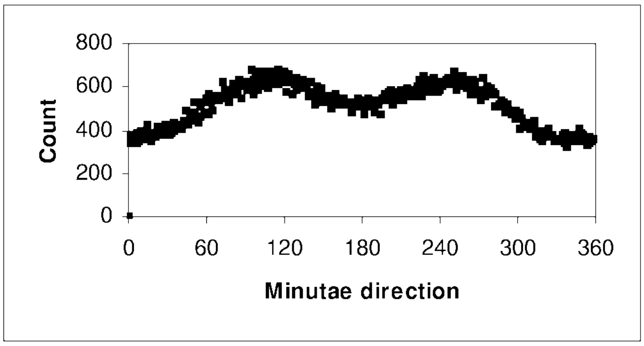
one set of minutia the differences between the aligned corresponding minutia coordinates and orientation angles were calculated. These difference represent attribute measurement errors and are presented in Fig. 1a–c.



(a)



(b)



(c)

Fig. 2. The prior distribution of minutiae attributes: (a), (b) x and y coordinates respectively.

It is clear that the distribution densities of all three attributes can be approximated by the normal distribution. The mean values and standard deviations for x, y coordinates and orientation angle are $(0, 5.1)$, $(0, 6.2)$ and $(-0.3, 6.6)$, respectively.

Note, that both images that are involved in the registration process have been preprocessed by the same feature extraction program. Consequently, the measurement errors should be distributed identically. This requires to use symmetric model for the estimate of measurement errors. Suppose the attributes of minutiae are estimated with some Gaussian error $x = x' + \xi_x$, $y = y' + \xi_y$, $\varphi = \varphi' + \xi_a$, where x, y, and φ are x, y coordinates and direction angle of minutiae, respectively. Let x', y', and φ' are the true values of the attributes and ξ_x, ξ_y, ξ_a are random measurement errors. We suppose that LSM registration aligns exactly the true position values of correctly matched minutia. Therefore the differences between the attributes of the matched minutia, for example, Δx can be expressed as

$$\Delta x = x_1 - x_2 = \xi_{x1} - \xi_{x2}.$$

Since ξ_{x1} and ξ_{x2} are independent and identically distributed random variables with mean and standard deviation equal to x' and σ_x, respectively, the mean and standard deviation of Δx are equal to 0 and $\sqrt{2}\sigma$. Consequently, in order to assess the standard deviation of x, y, and direction angle we should divide the experimentally obtained values 5.1, 6.2 and 6.6 by $\sqrt{2}$.

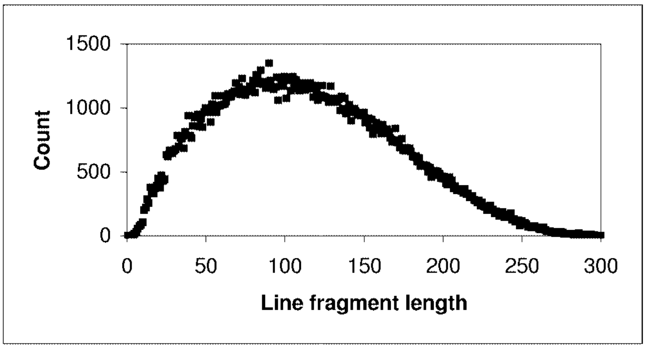
The next task was to estimate experimentally the prior distributions of attributes of the composite features. In particular we were interested in prior distributions of the invariants, that were used in feature matches filtering. In Fig. 2a–c, the histograms of x, y coordinates and direction angle of minutia, while in 3a–b the histograms of distances between the two minutia and angles between the x-axis and the line connecting the two minutia are presented. The distance and slope angle are the new attributes of composite features of cardinality two that where used in filtering process.

These histograms were obtained using minutiae features from 75 fingerprint images. As can be seen from Fig. 2, only the minutia orientation angle distributions are close to uniform, while the minutia position X and Y coordinates are distributed more close to the Gaussian distribution. This leads to the Chi-square distribution having two degrees of freedom for the distance between the minutia (Fig. 3). Nevertheless, the distribution of the distances between two minutia in the interval 50–150 is approximately uniform. The distribution of the line slope angles could also be approximated by uniform or truncated Gaussian distributions.

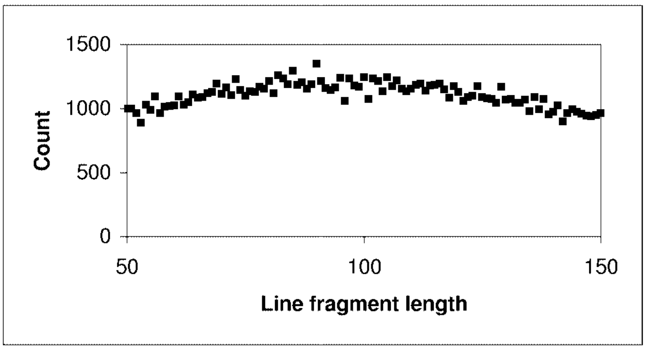
The distributions presented in this section supports the assumptions about the distributions of invariants made in Sections 5 and 6.

7.2. *Inlier/Outlier Ratio Increase*

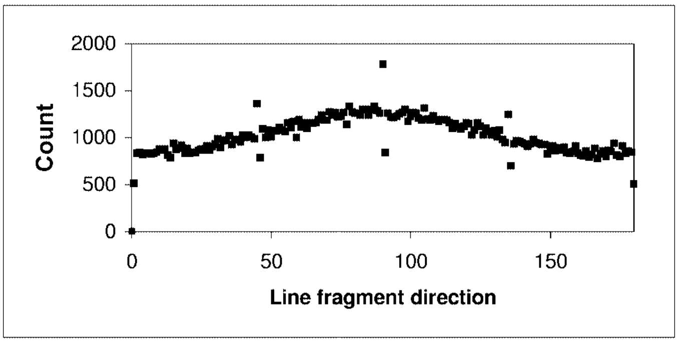
Having the manually established correct correspondences we were able to trace how the number of correct matches varies in comparison with the total number of the feature matches. All data were obtained using composite feature match filtering procedure from our previous paper (Malickas and Vitkus, 1999). This procedure was heuristic, but the



(a)



(b)



(c)

Fig. 3. The prior distribution of the composite feature (pair of minutia) atributes: (a), (b). Length of line fragment connecting two minutia in full range and in truncated range, respectively (c) line fragment slope angle.

parameters were selected according to the thorough experimental search and should be not very far from optimal ones.

In brief, we use the composite features consisting of two minutia and three invariant attributes were used to reject the false matches. The invariants were the distance between the minutia, and two differences between the direction angles of the two minutia and the slope angle of line connecting these minutia. The filtering procedure was as follows. First, we computed the differences between the corresponding invariants of two features of the hypothetical match. Then these differences are compared with two thresholds c_d, c_a (one for distance and one for angle differences) and the feature match was rejected if the difference was larger then the corresponding threshold. Finally, the euclidean norm of 3-dimensional vector of differences was calculated, and the match pair was rejected if the norm exceeds threshold $0.7^*(c_d + c_a)$. The following threshold values have been found experimentally to be optimal for $c_d = 13$, $c_a = 10$. In order to reduce the computational load, we restrict the distance interval to 50–150 and the total number of composite feature to 400. The starting point is the ratio ρ of correct and total numbers of basic features or minutia matches. In Fig. 4. the distribution of this initial ratio ρ is presented. This distribution was obtained from 50 pairs of the fingerprint images with manually established feature correspondences. As seen from the figure, the ratio does not exceed 3%.

As disclosed in (Malickas and Vitkus, 1999) feature consensus algorithm estimates transformation parameters in sequential manner. And the matches filtering could be performed before each stage using increasing number of invariants. The changes of the ratio of the correct and total match numbers in each stage are presented in Table 1. As expected, transition to composite features reduces the ratio to the level of less than 1%. The filtering process increases the ratio up to the range of 10–60%. After the rotation angle estimate, use of the additional invariant, the slope angle, increases the ratio to 20–90%, and full recovery of the transformation increases this ratio to 72% and more. In most cases the ratio exceeds 90%.

We also present the example when application of feature matches filtering process do not lead to the significant improvement of registration algorithm performance. Fig. 5 shows the improvement of Hough transform-based algorithm with the composite features of cardinality 2. Differently from the Hough type algorithm based on minutia features, the composite feature matches could be filtered using the same invariants, which were

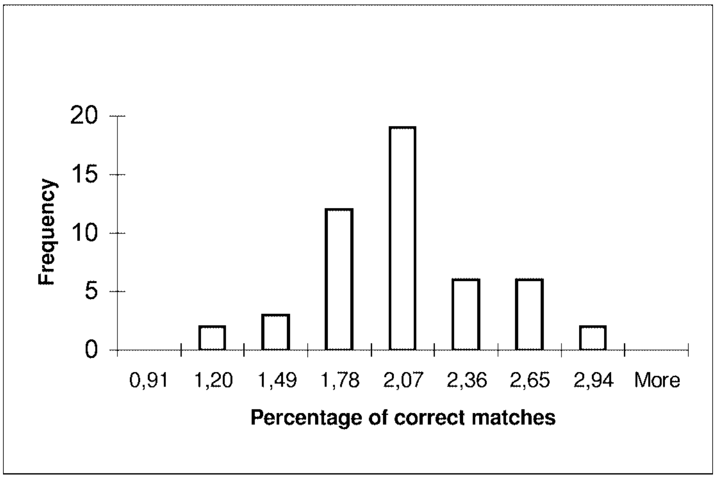


Fig. 4. Distribution of initial ratio of correct minutiae matches to total number of minutiae matches.

Table 1

Improvement of the ratio of correct matches number to total matches number in different stages of feature consensus algorithm

Initial		Rotation		Translation		Recognition	
Bin	Frequency	Bin	Frequency	Bin	Frequency	Bin	Frequency
0.026509	1	8.148148	1	19.04762	1	0	1
0.069553	30	12.8933	10	26.00951	8	9.090909	0
0.112598	43	17.63845	24	32.9714	6	18.18182	0
0.155643	25	22.38361	18	39.93329	20	27.27273	0
0.198687	1	27.12876	26	46.89518	16	36.36364	0
0.241732	4	31.87391	15	53.85707	22	45.45455	0
0.284777	15	36.61906	5	60.81896	17	54.54545	2
0.327821	4	41.36422	6	67.78085	13	63.63636	1
0.370866	0	46.10937	6	74.74274	13	72.72727	6
0.413911	0	50.85452	4	81.70463	6	81.81818	23
0.456955	0	55.59967	7	88.66652	2	90.90909	36
More	2	More	3	More	1	More	56

used in the first filtering step of the feature consensus algorithm presented in Table 1. Nevertheless, despite tenfold improvement of the ratio of correct matches number to total matches number, we notice only slight improvement of the ROC characteristic at the low level of false acceptance. While the transition from the minutia to minutia pair features in feature consensus algorithm gives about 10% improvement of the true acceptance rate at the same level of false acceptance rate (Malickas and Vitkus, 1999). This can be explained by better robustness of Hough transform-based algorithms in comparison to the feature consensus algorithm. The occupational model analysis conducted by Grimson *et al.* (Grimson, 1990) for Hough transform algorithms is also valid for feature consensus algorithn, since the latter uses similar voting scheme. According to this analysis, the probability that in a bucket of the accumulator array randomly will appear the peak of size l or larger is proportional to the redundancy factor, and inversely proportional to the total number of buckets in the accumulator array. Since the redundancy factor for both algorithms is the same, but the number of buckets in $3D$ Hough transform accumulator is much larger than in $1D$ feature consensus accumulator, the former is more robust to false matches.

8. Conclusions

Image registration and object recognition algorithms are mostly based on the establishment of local (basic) features correspondences. For each transformation group and basic feature type there is a minimum number of corresponding feature pairs (composite feature matches) necessary to recover uniquely the geometric transformation relating the two different images. So in order to recover transformation which correctly aligns the two im-

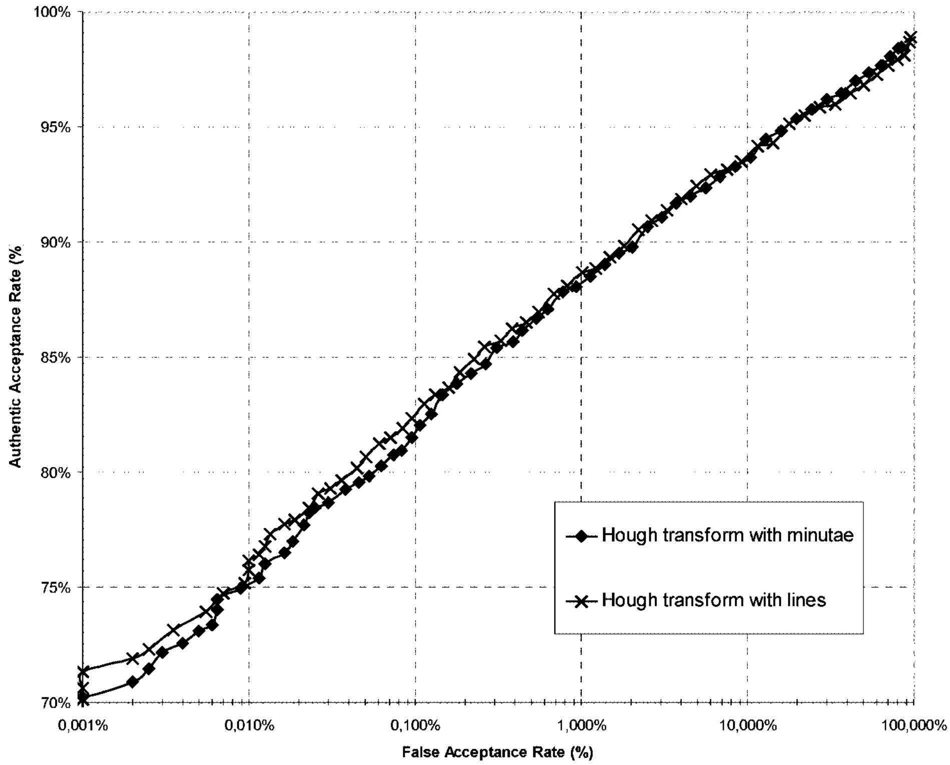


Fig. 5. Comparison of two Hough transform-based algorithms: a) based on minutia features without matches filtering and b) based on minutia pairs (lines) with feature matches filtering

ages we should have at least one correct minimal composite feature match. However, the correct matches are not known in advance, and many automatic image registration procedures have different algorithms to establish the correct correspondence between the features from different images. The main difficulty here is large number of false matches in comparison with the number of correct matches, and large percentage of features from the image that do not have corresponding counterparts in the other image, puts even more heavy burden on the task.

One of ways to reduce the number of hypothetical matches is to filter the feature matches using the feature attributes which are invariant under the action of transformations belonging to the considered geometric transformation group. If the minimum composite features are insufficient to generate invariants, they become available after increasing the cardinality of composite features by adding to them the additional basic features. While the feature match filtering has been widely used as a supplement measure to cope with large number of hypothetical matches, all these filtering techniques were heuristic. In this paper we proposed to treat feature matches filtering as an optimal statistical classification of matches into the correct or incorrect match classes. This enables to devise the optimized filtering process, or to evaluate possible improvements of suboptimal filtering procedures for a given invariant set.

In this paper we also surveyed two general frameworks allowing to incorporate the filtering procedure into the existing registration algorithm by increasing the cardinality of the composite feature. The examples of application of each framework are presented. Since the linear increase of the composite feature cardinality invokes exponential increase of hypothetical matches we have reviewed the published efficient implementation schemes, and proposed a new one.

All filtering implementations known to us use differences of the invariants as statistics for the incorrect match rejection. To construct the optimized classificator, it is necessary to estimate the conditional distribution densities or likelihood functions of this statistics. The simplest way to do that is to approximate the invariant distribution with the Gaussian distribution using well known linearizing technique under the assumption that the local feature attributes are measured with Gaussian error distribution. We demonstrate how it can be realized for invariants used in our own fingerprint images registration algorithm, and present experimental evidences indicating that general assumptions regarding feature distributions are valid in the case of the minutiae features from the fingerprint images.

Finally, the efficiency of suboptimal filtering process were presented, and it was demonstrated that final improvement of the overall registration process depends heavily on the robustness of the registration algorithm which follows the filtering stage.

References

Bergen J.R., and H. Shvaytser (1991). A probabilistic algorithm for computing Hough transforms. *Journal of Algorithms*, **12**, 639–656.

Costa M., R.M. Haralick and L.G. Shapiro (1990). Object recognition using optimal affine-invariant matching. Optimal affine-invariant point matching. In 6^{th} *Israel Conf, Art. Intell.*, pp. 35–61.

Can A., C.V. Stewart and B. Roysam (1999). Robust hierarchical algorithm for constructing a mosaic from images of the curved human retina. In *Proc CVPR*, pp. 286–292.

Fischler M.A., and Bolles R.C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, **24**, 381–396.

Fomin J.A., G.R. Tarlovskij (1986). Statistical theory of pattern recognition. *Radio and communications*, Moscow (in Russian).

Gold S., A. Rangarajan, C. Lu, S. Pappu and E. Mjolsness (1998). New algorithms for $2D$ and $3D$ point matching: pose estimation and correspondence, *Pattern Recognition*, **31**(8), 1019–1031.

Govindu V., C. Shekhar (1999). Alignment using distributions of local geometric properties. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **21**(10).

Grimson W.E., and D.P. Huttenlocher (1988). On the sensitivity of the Hough transform for object recognition, IEEE trans. *PAMI*, **12**(3), 255–274; Also in *MIT AI lab Memo*, **1044**, 1988.

Grimson W.E., D.P. Huttenlocher, D.W. Jacobs (1991). Affine matching with bounded sensor error: a study of geometric hashing and alignment. XXX; Also in *MIT AI lab Memo*, **1250**.

Grimson W.E., D.P. Huttenlocher and T.D. Alter (1992). Recognizing $3D$ objects from $2D$ Images: an error analysis., *Proc. IEEE conf. On Computer vision and pattern recognition*; also *MIT AI Lab Memo* **1362**.

Hong L. *et al.* (1998). Fingerprint image enhancement: algorithms and performance evaluation. *IEEE Trans. Pattern Analysis and Machine Inteligence*, **20**(8), 777–789.

Hsieh J.W, H.Y. Mark Liao, K.C. Fan, M-T. Ko and Y.P. Hung (1997). Image registration using a new edge-based approach. *CVGIP: Computer Vision and Image Understanding*, **67**(2), 112–130.

Hummel R. and H. Wolfson (1988). *Affine Invariant Matching*. DARPA Image Understanding workshop.

Huttenlocher D.P., S. Ullman (1987). Recognizing rigid objects by aligning them with an image. *MIT AI lab Memo*, **937**.

Jacobs D.W. (1989). Grouping for recognition. *MIT AI Memo*, **1177**.

Lambdan Y.J., J. Schwartz and H. Wolfson (1988). Object recognition by affine invariant matching. In *IEEE Computer Vision and Pattern Recognition Conf.*, pp. 335–344.

Lambdan Y.J. and H. Wolfson (1988). Geometric hashing: a general and efficient model based recognition scheme. In *Second International Conference on Computer Vision.*, pp. 238–249.

Lester H., Simmon R. (1999). A survey of Hierarchical non-linear medical image registration. *Pattern Recognition*, **32**, 129–149.

Levin B.R. (1989). Theoretical foundations of statistical radiotechniques. *Radio and Communications*, Moscow, p. 653 (in Russian).

Lowe D.G. (1987). Three-dimensional object recognition from single two–dimensional images. *Artificial Intelligence*, **31**(3), 355–395.

Malickas A., and Vitkus R. (1999). Fingerprint registration using composite features consensus. *Informatica*, **10**(4), 389–402.

Olson C.F. (1992). Fast alignment by elimination unlikely matches. Univ. of California at Berkley. Technical Report UCB//CSD-92–704.

Olson C.F. (1997). Efficient pose clustering using a randomized algorithm. *Int. J. of Computer Vision*, **23**(2), 131–147.

Olson C.F. (2000). A general method for feature matching and model extraction. In *Vision Algorithms: Theory and Practice*, 20–36.

Ratha N., Karu K., Chen S., A. Jain (1996). A real time matching system for large fingerprint database. *IEEE Trans. Pattern Analysis and Machine Inteligence*, **13**(8), 799–813.

Shechar Ch., V. Govindu, R. Chellappa (1999). Multisensor image registration by features consensus. *Pattern Recognition*, **32**, 39–52.

Tsai F. (1993). A probabilistic approach to geometric hashing using line features. *Technical Report 640*, Robotics research laboratory, N.Y.U., June.

Tsai F.C. 621. Statistical approach to affine invariant matching with line features, N.Y. University, Robotics research lab. TR-621.

Weinshall D. (1993). Model-based invariants for $3D$ vision. *Intern. J. of Computer Vision*, **10**(1), 27–42.

Wells W.M. (1993). Statistical object recognition. *MIT AI Technical report*, **1398**.

Williamson T.A. (1998). *A High Performance Stereo Vision System for Obstacle Detection*. Carnegie Mellon Univ. Robotics institute, TR 98–24.

Xu L., E. Oja, and P. Kultanen (1990). A new curve detection method: Randomized Hough transform (RHT). *Pattern Recognition Letters*, **11**, 331–338.

Zheng Q., R. Chellappa (1995). Automatic feature point extraction and tracking in image sequences for arbitrary camera motion. *I,J, Computer Vision*, **15**, 31–76.

A. Malickas has graduated from Vilnius University (Faculty of Physics) in 1989. His research interests include biometrical person identification, image processing and recognition, neural networks.

R. Vitkus has graduated from Vilnius University (Faculty of Nature Science) in 1980. He receive Ph.D. degree from the Moscow Institute of Information Transmission Problems. His research interests include image processing and recognition, biometrical person identification.

Požymių lyginimų filtravimas pagal geometrinius invariantus vaizdų registravime

Algimantas MALICKAS, Rimantas VITKUS

Požymių lyginimų filtravimas, plačiai naudojamas vaizdų registravime, yra euristinis metodas skirtas sumažinti požymių palyginimų skaičiui. Darbe siūloma tokį lyginimo filtravimą interpretuoti kaip klasifikacijos, į teisingus ir neteisingus lyginimus, uždavinį. Statistika, pagal kurią toks klasifikavimas atliekamas, yra paremta elementarių požymių sutvarkytų rinkinių (sudėtinių požymių) invariantų skirtumais. Siūloma tokio klasifikavimo efektyvaus įdiegimo schema ir optimizavimo būdai. Pateikti eksperimentiniai rezultatai, iliustruojantys siūlomų metodų efektyvumą.