SHAPE CLASSIFICATION OF PARTIALLY OCCLUDED OBJECTS USING SUBSPACE DETECTORS

Arnt-Børre Salberg[a], Alf Harbitz[a], and Alfred Hanssen[b]*

[a] Institute of Marine Research, Dept. Tromsø, NO-9294 Tromsø, Norway
[b] University of Tromsø, Dept. of Physics, NO-9037 Tromsø, Norway

ABSTRACT

In this paper we consider shape classification of partially occluded objects. We model the occlusion as non-Gaussian noise, and apply robust subspace detectors in the classification module. We show that the robust subspace detectors can be formulated as a weighted subspace detector, and the elements in the boundary vector will be weighted before they are matched. The part of the boundary vector that corresponds to the occluded part of the object, will be suppressed by the weight vector, and hence have a reduced effect on the classification performance. The detectors are demonstrated on fish species classification applications.

1. INTRODUCTION

The classification of a two-dimensional (2-D) occluded object from its boundary curve represents a challenging and important problem in several applications (e.g. [1]). In many real-world images the objects of interest can be viewed approximately as 2-D.

Kendall [2] defines the shape of an object as "what is left when the differences which can be attributed to translations, rotations, and dilations have been quotiented out". We represent a 2-D shape by a set of complex samples obtained from the contour of the object. These complex samples constitute a vector which represents objects as points on a high-dimensional manifold, called the shape space [2, 1], or as a 1-D subspace of $\mathbb{C}^N$. Object recognition can then be achieved by a subspace detector [3, 4, 5], which possesses a natural invariance to rotations and scalings of the object.

We show that by modeling the occlusion as non-Gaussian (and heavy-tailed) noise, the object recognition rate increases by using robust space detectors (e.g. as proposed in [5]). In this paper we show that we may recast the robust subspace detectors proposed in [5] as a weighted subspace detector, that is, the measurements are weighted before they are matched to an oblique projection of the signal. Furthermore, we may base the subspace detector on the Least-Median-of-Squares (LMedS) estimator [6], which leads to a binary weighting matrix (i.e. the elements are zero or one). This weighted subspace detector can be interpreted as a subspace detector with interference on the elements corresponding to a zero weight factor.

The goal is two-fold: 1) To extend the existing subspace detection schemes, and propose robust methods for classification of partially occluded objects, and 2) illustrate the proposed methods to an example of fish species classification.

We will denote matrices and vectors in bold face, using capital letters for matrices and lower case letters for vectors. The transpose, conjugate transpose, and the Moore-Penrose pseudo-inverse of a matrix $\mathbf{A}$ will be denoted by $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^\dagger$, respectively. $[\mathbf{A}]_{m,n}$ will denote its (m,n)th element, $\langle\mathbf{A}\rangle$ the subspace spanned by its columns, and $\mathbf{P}_A = \mathbf{A}\mathbf{A}^\dagger$, the orthogonal projection matrix onto the subspace $\langle\mathbf{A}\rangle$. A linear operator $\mathbf{E}$ is a projection if and only if it is idempotent, i.e. $\mathbf{E} = \mathbf{E}^2$.

2. LANDMARK VECTOR EXTRACTION

To construct the N dimensional landmark vector $\mathbf{x}'$ from the binary image, some normalization procedure is necessary in order to be able to associate the contour to the corresponding model in recognition applications.

First, N_k initial complex boundary points $\mathbf{q}_k = [q_1, \ldots, q_{N_k}]^T$ are extracted from the binary image, and normalized with respect to its center-of-gravity. To reduce the effect of deformations on the extracted boundary, a simple μ-law companding operator (see e.g. [7]) is applied on the radius of each 2-D boundary point. Since large deformations is suppressed by the μ-law compandor, the principal axis is calculated from the μ-law compressed shape. Furthermore, the compressed shape is normalized with respect to its center-of-gravity, and the inverse companding operator is applied. In order to ensure that all shapes are represented by a fixed (and equal) number of samples, the elements in $\mathbf{q}_k$ nearest to the points where the major principal axis intersects the object contour are detected. Then, $N/2$ equispaced points are extracted between the two intersect

*Thanks to The Research Council of Norway (NFR) for funding under project 052703.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

points. These N points, which are stacked into the vector (often called landmark vector) $\mathbf{x}' = [z_0, z_1, \ldots, z_{N-1}]^T$, $z_n = x_n + iy_n$, represents the boundary of the object.

3. THE SHAPE SPACE CONCEPT

The shape space concept was introduced by Kendall in [2], and in this section we review some relevant aspects of the 2-D shape space theory.

To remove the effect of translation, we consider the vector

$$\mathbf{x} = [z_0 - \overline{z}, z_1 - \overline{z}, \ldots, z_{N-1} - \overline{z}]^T, \tag{1}$$

where $\overline{z} = (1/N)\sum_{n=0}^{N-1} z_n$ is the centroid of $z_0, z_1, \ldots, z_{N-1}$. Since $\sum_{n=0}^{N-1}(z_n - \overline{z}) = 0$, the point $\mathbf{x}$ is located on a $N-1$ dimensional complex hyperplane passing through the origin of $\mathbb{C}^N$. To remove the effect of scaling and rotation, we model the boundary vector as

$$\mathbf{x}^* = \lambda\mathbf{x}, \quad \lambda \in \mathbb{C}\backslash\{0 + j0\}. \tag{2}$$

The complex number λ covers all possible scalings and rotations of $\mathbf{x}$, and hence the vector $\mathbf{x}^*$ represents the shape of $\mathbf{x}'$. Note, that since the boundary vector $\mathbf{x}'$ is founded on the principal intersect points, there is an ambiguity among the elements in the boundary vector. For fish this amounts to which of the two principal intersection points that belongs to the head and tail of the fish. This ambiguity can be represented by a cyclic permutation of the elements of the boundary vector, i.e., we model the boundary vector as

$$\mathbf{x}^{**} = \lambda\mathbf{C}^n\mathbf{x}, \quad n \in \{0, N/2\} \tag{3}$$

where $\mathbf{C}$ cyclically permutes the elements on position downwards, and $n \in \{0, N/2\}$ denotes the order of the permutation.

The point $\mathbf{x}^{**}$ represents a point in the complex projective space $\mathbb{CP}^{N-2}$ [2]. This is a smooth and curved non-Euclidean space which is called *shape space*. We will not go into the details about the shape space, but base our derivations on a (simple) linear model. For a comprehensive discussion on shape space theory we refer the reader to [8].

4. SUBSPACE DETECTORS

Let the observed boundary vector of the mth object be modelled as

$$\mathbf{y} = \lambda\mathbf{C}^n\mathbf{x}_m + \mathbf{w}, \tag{4}$$

where $\mathbf{w}$ denotes an additive zero-mean white complex Gaussian noise vector with covariance matrix $\sigma^2\mathbf{I}$, and the scaling/rotation parameter λ and the order of the permutation n are unknown.

The detector is derived using the Generalized Likelihood Ratio (GLR) principle (see e.g. [3, 4]). The GLR replaces the unknown parameters (in this case λ and n) by their Maximum Likelihood Estimates (MLE). The MLE of λ is [3]

$$\widehat{\lambda} = (\mathbf{C}^n\mathbf{x}_m)^\dagger\mathbf{y}, \tag{5}$$

Note that it is only for 2-D signals that we may use the unconstrained MLE estimates, since a complex scalar does not introduce any affine transformation of the shape.

Now, deleting all class independent terms, the resulting GLR detector can be written as

$$J_m = \max_{n\in\{0,N/2\}} \mathbf{y}^H\mathbf{P}_{C^n x_m}\mathbf{y} = \max_{n\in\{0,N/2\}} \|\mathbf{P}_{x_m}\mathbf{C}^{-n}\mathbf{y}\|_2^2. \tag{6}$$

The detector first compensates for the effect of cyclic permutation by applying $\mathbf{C}^{-n}$ on $\mathbf{y}$, and then it projects the signal $\mathbf{C}^{-n}\mathbf{y}$ onto the shape subspace $\langle\mathbf{x}_m\rangle$. The decision is then based on the amount of energy of $\mathbf{C}^{-n}\mathbf{y}$ that resides in the subspace $\langle\mathbf{x}_m\rangle$. Note that for a given n, this is equivalent with the geodesic distance detector [2, 1]

4.1. Object class recognition

The subspace detector in Eq. (6) can easily be extended to an object class detector, e.g., classify a boundary vector $\mathbf{y}$ to be a fish, ship, or plane.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_K$ denote a set of boundary templates that all belong to the same class (e.g. a fish), and construct the matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_K]$. Now, an object class detector can be formulated as

$$J = \max_{n\in\{0,N/2\}} \mathbf{y}^H\mathbf{P}_{C^n X}\mathbf{y} = \max_{n\in\{0,N/2\}} \|\mathbf{P}_X\mathbf{C}^{-n}\mathbf{y}\|_2^2. \tag{7}$$

Note that this detector is similar to the one in Eq. (6), but now we project $\mathbf{C}^{-n}\mathbf{y}$ onto the K-dimensional shape subspace $\langle\mathbf{X}\rangle$ that accounts for all shape variations within an object class.

5. SHAPE CLASSIFICATION OF PARTIALLY OCCLUDED OBJECTS

Occlusion induces object shape distortions, in the sense that some of the object boundary points are false. Moreover, an occlusion also shifts the center-of-mass of the object, and changes the directions of the principal axes. In practice, one may not know which boundary points that are “true”, and hence an objective and automatic procedure to suppress the effect of occlusion is needed.

If the noise vector is non-Gaussian, or the interference subspace is an arbitrary unknown subspace, we may consider to use some of the robust detectors proposed in e.g. [5]. The idea is to suppress the effect of outliers in $\mathbf{y}$, and an often used method is o consider the p-norm instead of the squared Euclidean distance. Thus, we may represent the discriminant function as [5]

$$J_m = \|\mathbf{y} - \mathbf{C}^n \mathbf{x}_m \widehat{\lambda} - \mathbf{1}\widehat{v}\|_p^p, \tag{8}$$

where $\widehat{\lambda}$ and $\widehat{v}$ are the estimates of λ and $\mathbf{v}$, respectively, and $\|\mathbf{r}\|_p$ for a given vector $\mathbf{r}$ is defined as

$$\|\mathbf{r}\|_p^p = \sum_i |r_i|^p. \tag{9}$$

While many robust estimation techniques have been proposed in the statistics literature, the two main ones used are M-estimators and least-median-of-squares (LMedS). We will now design subspace detectors using these techniques.

5.1. Subspace detectors based on M-estimators

To calculate the M-estimate $\widehat{\lambda}$ and $\widehat{v}$ for $p \neq 2$, we often apply an iterative procedure, where we solve the following iterated re-weighted least-squares problem (see e.g. [6])

$$\begin{bmatrix} \widehat{\lambda}^{(k)} \\ \widehat{v}^{(k)} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{x}_m^H \mathbf{C}^{-n} \\ \mathbf{1}^H \end{bmatrix} \mathbf{W}^{(k-1)} [\mathbf{C}^n \mathbf{x}_m, \mathbf{1}] \right)^{-1} \cdot \begin{bmatrix} \mathbf{x}_m^H \mathbf{C}^{n} \\ \mathbf{1}^H \end{bmatrix} \mathbf{W}^{(k-1)} \mathbf{y}, \tag{10}$$

where $\mathbf{W}^{(k-1)} = \operatorname{diag}\{w(r_0^{(k-1)}), \ldots, w(r_{N-1}^{(k-1)})\}$, $r_i^{(k-1)}$ is the ith element of $\mathbf{r}^{(k-1)} = \mathbf{y} - \mathbf{C}^n \mathbf{x}_m \widehat{\lambda}^{(k-1)} - \mathbf{1}\widehat{v}^{(k-1)}$, and the superscript (k) indicates the iteration. The weight $w(r_i^{(k-1)}) = |r_i^{(k-1)}|^{p-2}$ should be recomputed after each iteration number in order to be used in the next iteration. The discriminant can in this case be expressed as [9]

$$J_m^{(k-1)} = \min_{n \in \{0, N/2\}} \mathbf{y}^H (\mathbf{W}^{(k-1)} - \mathbf{W}^{(k-1)} \mathbf{E}^{(k-1)}) \mathbf{y}, \tag{11}$$

where

$$\mathbf{E}^{(k-1)} = [\mathbf{C}^n \mathbf{x}_m, \mathbf{1}] \left(\begin{bmatrix} \mathbf{x}_m^H \mathbf{C}^{-n} \\ \mathbf{1}^H \end{bmatrix} \mathbf{W}^{(k-1)} [\mathbf{C}^n \mathbf{x}_m, \mathbf{1}] \right)^{-1} \cdot \begin{bmatrix} \mathbf{x}_m^H \mathbf{C}^{n} \\ \mathbf{1}^H \end{bmatrix} \mathbf{W}^{(k-1)} \tag{12}$$

is an oblique projection onto the subspace $\langle [\mathbf{C}^n \mathbf{x}_m, \mathbf{1}] \rangle$ The detector works as follows: First, the measurement vector is weighted by $\mathbf{W}^{(k-1)}$ to suppress large amplitudes in $\mathbf{y}$, and then matched to its oblique projection $\mathbf{E}^{(k-1)}\mathbf{y}$. The iterations continue according to some stopping rule, e.g. until the absolute value of the relative difference of two succeeding J_m-values is less than a given threshold.

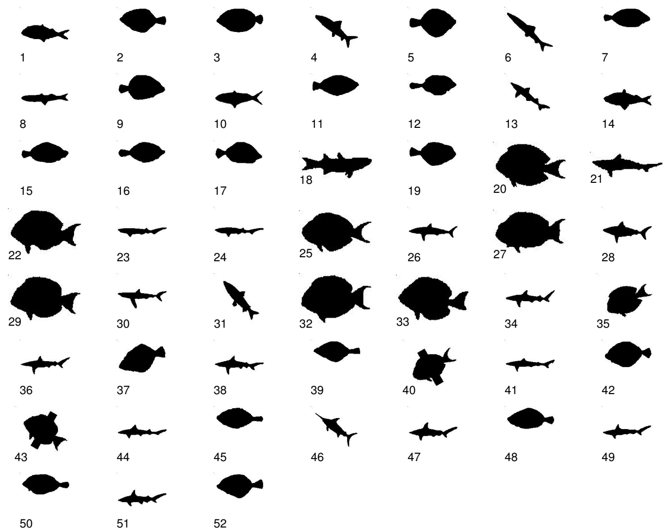


Fig. 1. The fish template database.

5.2. Subspace detectors based on LMedS

Another robust parameter estimation technique is the method of least-median-of-squares, which aims to solve the non-linear minimization problem [6]

$$\{\widehat{\lambda}, \widehat{v}\} = \arg\min_{\lambda, v} \operatorname*{med}_i r_i^2. \tag{13}$$

That is, the estimator must yield the smallest value for the median of squared residuals [6]. Unlike the robust detectors based on the M-estimate, the LMedS problem cannot be solved using re-iterated weighted least squares. The idea is to randomly select a number of C-point subsets of the N data points. A parameter vector $[\lambda_j, v_j]$ is fitted to the points in each subset using e.g. least squares. In addition, the median of the squared all remaining residuals $r_i^2(\lambda_j, v_j)$, denoted by M_j, is also determined by

$$M_j = \operatorname*{med}_i r_i^2(\lambda_j, v_j). \tag{14}$$

The LMedS solution are the estimators for which the corresponding M_j is the minimum among all different M_j's.

The LMedS procedure can now be used to determine a set of binary weights $w_i \in \{0, 1\}$, on which we can build our subspace detector (and which will modify the estimators of λ and v [6]). We now have that

$$\{\widehat{\lambda}, \widehat{v}\} = \arg\min_{\lambda, v} \sum_i w_i r_i^2, \tag{15}$$

where the weight w_i is [6]

$$w_i = \begin{cases} 1, & \text{if } r_i^2 \leq (2.5\widehat{\sigma})^2, \\ 0, & \text{otherwise,} \end{cases} \tag{16}$$

and $\widehat{\sigma}$ is the robust standard deviation given as [6]

$$\widehat{\sigma} = 0.8493[1 + 5/(N(1-\epsilon))]\sqrt{M_{min}}, \qquad (17)$$

where M_{min} is the minimum median

$$M_{min} = \min_{j} \operatorname*{med}_{i=1,\ldots,N} r_i^2(\widehat{\lambda}_j, \widehat{v}_j), \qquad (18)$$

and ϵ is the maximum fraction of outliers in a measurement vector. From the derivation above we see that the core quantity in the LMedS algorithm is the robust standard deviation $\widehat{\sigma}$, which determines the elements that is likely to include noise outliers.

In the LMedS case, the weight matrix $\mathbf{W}$ is a projection matrix since $\mathbf{W}^2 = \mathbf{W} \triangleq \mathbf{P}_W$, and the discriminant function can be formulated as the following subspace detector

$$J_m = \min_{n\in\{0,N/2\}} \mathbf{y}^H(\mathbf{P}_W - \mathbf{P}_{W[C^n x_m, S]})\mathbf{y} \qquad (19)$$

Note that this detector corresponds to the signal model $\mathbf{y} = \mathbf{C}^n\mathbf{x}_m\lambda + \mathbf{1}v + \mathbf{N}_I\mathbf{u} + \mathbf{w}$, where $\mathbf{N}_I$ models the occlusion subspace. The columns of $\mathbf{N}_I$ are selected from the columns of the identity matrix that corresponds to diagonal elements equal to zero in the weight matrix $\mathbf{W}$, and the occlusion amplitudes are determined by $\mathbf{u}$. Clearly, $\mathbf{P}_{N_I} = \mathbf{P}_W = \mathbf{W}$.

6. NUMERICAL EXAMPLE

In this section we provide an example of the robust subspace detectors presented in Eqs. (11) and (19).

Fig. 1 shows the set of 52 fish shapes that constitute the templates we applied in the numerical example. Each shape was represented by $N = 64$ boundary points. For each of the detectors we applied a rotated an occluded object (see the first column in Table 1) to match against each template shown in Fig. 1. The true object is object number 4 in Fig. 1. The rotation angle was selected systematically from -180 to 168.75 degrees (with a $11.25°$ step). For the detector in Eq. (11) the weight was chosen to be $w(r) = |r|^{-1}$ (corresponds to $p = 1$ in Eq. (8)), and for the LMedS based detector in Eq. (19) we chose $\epsilon = 0.15$, $C = 8$, and the number of C-point subsets equal to 100. For both detectors the occlusion subspace matrix was $\mathbf{S} = \mathbf{1}$.

The recognition rates of the detectors applied to the example described above are shown in Table 1. For both objects, the LMedS based subspace detector yields the best performance (88% and 31% correct classifications). The subspace detector based on M-estimates achieved 88% and 0% correct classifications.

7. CONCLUSIONS

We have proposed a shape object recognition method using robust subspace detectors. The robust detectors treated occlusion as outliers in the boundary vector, and were based on iterated re-weighted least squares. The method were demonstrated on fish species classification, and outperformed the non-robust subspace detector.

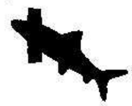
	Eq. (6)	Eq. (11)	Eq. (19)
	44%	88%	88%
	0%	0%	31%

Table 1. Recognition rates for the subspace detectors. The images show the occluded fish images that are being tested.

8. REFERENCES

[1] J. Zhang, X. Zhang, H. Krim, and G. G. Walter, "Object representation and recognition in shape spaces," *Pattern Recognition*, vol. 36, pp. 1143–1154, 2003.

[2] D. G. Kendall, "Shape manifold, procrustean metrics, and complex projective space," *Bull. London Math. Soc.*, vol. 16, pp. 81–121, 1984.

[3] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 2146–2157, 1994.

[4] A. B. Salberg and A. Hanssen, "Stochastic multipulse-PAM: A subspace modulation technique with diversity," *Signal Processing*, vol. 83, no. 12, pp. 2559–2577, Dec. 2003.

[5] M. N. Desai and R. S. Mangoubi, "Robust Gaussian and non-Gaussian matched subspace detection," *IEEE Trans. Signal Processing*, vol. 51, no. 12, pp. 3115–3127, Dec. 2003.

[6] Z. Zhang, "Parameter estimation techniques: A tutorial with applications to conic fitting," *Image Vision Computing*, vol. 15, pp. 59–76, 1997.

[7] J. D. Gibson, *Digital and Analog Communications*, Prentice-Hall, Upper Saddle River, N.J., 1993.

[8] D. G. Kendall, D. Barden, T. K. Carne, and H. Le, *Shape and Shape Theory*, Wiley, Chichester, England, 1999.

[9] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, Upper Saddle River, N.J., 2000.