CVC Tech.Rep. #035 October,1999

Real Time Recognition of Pharmaceutical Products by Subspace Methods

Jordi Vitrià, Petia Radeva, Xavier Binefa, Albert Pujol,
Ernest Valveny, Robert Benavente, Craig von Land
Centre de Visió per Computador - Departament
d'Informàtica
Edifici O, Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona) Spain

Real Time Recognition of Pharmaceutical Products by Subspace Methods

Jordi Vitrià, Petia Radeva, Xavier Binefa, Albert Pujol,
Ernest Valveny, Robert Benavente, Craig von Land
Centre de Visió per Computador - Departament d'Informàtica
Edifici O, Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona) Spain

Abstract

A prototype vision system has been developed for indexing pharmaceutical products. The system consists of a lighting and imaging dome, conveyor, color camera, and computer for image acquisition and analysis. Objects are represented by four indexes: size, color, edges and appearance. Indexes are computed by projecting original features in an optimal representation subspace resulting from Principal Component Analysis. The body of this paper details the design and operation of the system. Also presented are the results of an empirical study in which the system has been extensively tested in the indexing of 7458 different products. This study shows that an integrated color-, shape- and edge-based feature representation results in 99.05 % of the images being retrieved within the top five positions.

Key words: Subspace methods - Object recognition - Indexing - Appearance-based representation.

1 Introduction

Object indexing and recognition has been an active field of research in computer vision in the last years. Aspect-based methods have appeared as a powerful alternative to traditional 3D geometry based techniques ([Lowe 1987, Huttenlocher and Ullman 1987, Grimson 1990]) when geometrical models of the viewed objects can be difficult to obtain. Some problems remain difficult to be solved, like partial object occlusions and severe lighting changes, but important applications have been solved using these techniques.

Turk and Pentland (1991) used subspace methods to describe face patterns with a lower-dimensional space than the image space. The appearance of a face is a combination of its shape, reflectance properties, pose in the scene and illumination conditions. All these factors can be compactly represented using a well-known compression technique: principal component analysis (PCA), also known as the Karhunen-Love transform. It uses the eigenvectors of an image set for representing each image in the set. PCA can be used for dimensionality reduction, yielding projection directions that maximize the total scatter across all images. The problem of general object recognition and pose estimation was tackled by [Murase and Nayar (1993)]. They represent objects as manifolds in a low dimensional subspace formed by the n dominant eigenvectors of a set of training images representing all possible object views under all possible illumination conditions. Recognition is achieved by finding the manifold that is closest to the projection of an input image in the eigenspace formed by all objects. [Black and Jepson (1996)] have addressed the problem of partial occlusion by using robust estimation techniques. [Buhman et al. (1990)] use *local* appearance-based models, based on Gabor filter responses which are used in a graph matching strategy for recognition, for minimizing object deformations. The responses of a set of local "feature" detectors have also been used in different forms by several authors ([Martinez and Vitrià 1997, Schielle and Crowley 1996, Rao and Ballard 1995]).

Color distributions can be efficiently used as signatures for object recognition in the appearance-based framework. The earliest approach to this problem (Swain and Ballard, 1991) showed the usefulness of color histograms for indexing large object databases independently of object's pose. Most of the recent approaches focus on illumination color invariance (Funt and Finlayson, 1995, Gevers and Smeulders, 1997) known as color constancy, but although these methods perform better than histogram indexing when color illumination changes, they use color information only where surface color varies and are very sensitive to noise.

Appearance-based methods have only been tested in very-large image databases for the face recognition problem ([Phillips et al. 1998]), resulting in recognition rates around 85-90 per cent. Color based indexing has also been tested in very large image databases, being an illustrative example that referred in Gevers and Smeulders (1997).

The system presented in this paper integrates appearance based recognition and color indexing in a complete and operational prototype to recognize pharmaceutical product boxes. This problem constitutes a large scale test for these techniques, given that in our case (legal pharmaceutical products in Spain) this involves more than 10000 different products. Among these, we have identified 3500 "non-processable" products, corresponding to different causes (highly reflective box surfaces, plastic bags, very large objects, etc.), which have not been considered in the test. Figure 1 shows some accepted products, and figure 2 shows some rejected products.

2 System description

The system consists of three major components: the image acquisition component, the object localization component, and the object indexing component. This section describes each of these components and their integration.

Figure 1: Examples of considered pharmaceutical products

Figure 2: Examples of rejected pharmaceutical products

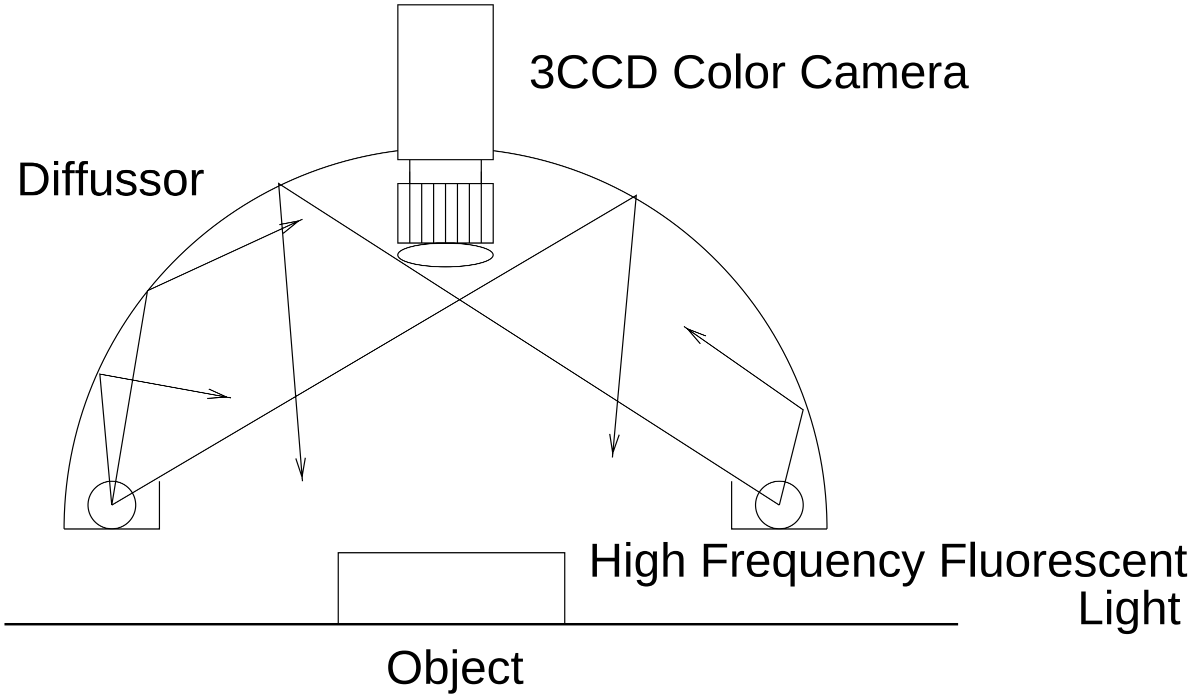

Figure 3: Lighting and imaging

2.1 Image Acquisition

For the capture of the image, a high resolution 3CCD color camera is employed. The camera is coupled with a standard frame grabber with a resolution of 768×494 pixels. The optics of the camera consists of a 25mm lens. This parameter allows for the acquisition of images corresponding to objects ranging from 2cm $\times$ 2cm to 20cm $\times$ 25cm. Due to the fact that products are transported by a continuous moving conveyor belt, only one field of the captured image can be considered. Hence, the useful pixel resolution is 384×247. Objects are placed on the conveyor belt by a human operator that assures that the most representative face of each box is presented to the camera [1].

In order to create a learning set of images, six images of every product were acquired, following a placement protocol to assure different object orientations with respect to the optical axis and different object locations in the field of view.

Lighting deserves special attention. In order to minimize lighting effects in the acquisition process, we have designed a lighting dome which is shown in figure 3. In spite of the fact that this lighting architecture minimizes object aspect changes due its relative position, it does not eliminate them, as can be seen in the image set shown in figure 4. This fact was decisive for deciding the acquisition of six images for each product.

2.2 Object localization and normalization

The first processing step when acquiring an image is the localization of the object in the image, performing some kind of figure-ground segmentation. This is not a serious problem in our case, since the

[1] The most representative face of the product corresponds to the one where commercial information -name, company, etc.- is printed.

Figure 4: Lighting effects due to highly reflective surfaces

conveyor was chosen to be non reflective black, and objects can be segmented by a constant threshold value. Exact object localization can be found following these steps:

1. We apply a constant threshold value to the acquired image, resulting in a set of blobs which are not necessarily connected. See figure 5 for an illustrative example.

2. The edges of the blobs are detected and stored in an efficient data structure. A modified Hough transform is used in order to detect the larger and most probable rectangular object of the image.

The aim of the segmentation process is to obtain the orientation, position and extent of the product box. We have handled this problem through the straight lines Hough transform. It has been chosen in order to avoid the problem related to boxes that, due to their colors, present a contour with holes or non connected regions (figure 6). Each contour point of the box accumulates a sinusoidal curve in the Hough space, where each point of the sinusoidal is a vote for one of the possible straight lines (defined in a polar representation, angle and distance to the origin) that passes through that point. The method is designed to recognize rectangular boxes. This means that the four lines that describe the boundary of the box will be disposed in the accumulation space so that, the two highest maxima (the most voted lines), corresponding to the longer box side, will appear in the same column. The maxima corresponding to the shorter side will appear in a column shifted 90 degrees with respect to the previous one. Once all the points of the boundary of the box have been accumulated, the box orientation, size and position are obtained using the position of the maxima. Thus, the position of the global maximum gives us the box main orientation and the distance between the two maxima of this column give us the size of the shortest side of the box. The distance between the most distant maxima in the column of the shortest side (column shifted 90 degrees respect to the previous one), gives the size of the longest box axis.

Once the box has been located, its image is warped to a normalized box size of 32x32 pixels. This normalization ensures a common image representation for all boxes, a necessary condition to perform PCA of a set of images (see section 2.3). The normalization method is unable to distinguish between two rectangular boxes rotated 180 degrees. To avoid this problem a 180 degrees rotational invariant transformation of the normalized image has been applied. The transformation applied works as follow:

Given the point (x, y), we will call $g(x, y)$ the position of this point when a 180 degrees rotation is applied. It is,

$$g(x, y) = (Tx - x, Ty - y)$$

where Tx, Ty is the size of the image (in our case $Tx = 32$ and $Ty = 32$ pixels). Then the original image $I(x, y)$ is transformed to its invariant representation $F[I(x, y)]$, where,

$$F(I(x, y)) = \begin{cases} |I(x, y) - I(g(x, y))| & \text{if } y < 16 \\ \frac{(I(x,y)+I(g(x,y))}{2} & \text{if } y \geq 16 \end{cases}$$

It can be seen easily that this transformation is invariant to 180 degrees rotations, that is,

$$F(I(x, y) = F(I(g(x, y))$$

A small amount of the considered boxes are squared, and in these cases, instead of rotations to 180 degrees we have to take into account their rotations to 0, 90, 180 and 270 degrees. In these cases the image of the box and its 90 degrees rotation are considered as two independent descriptors of the same product. This duplication of descriptors does not represent an important decrease of the efficiency of the system due to the reduced number of products with exactly squared boxes.

2.3 Object indexing

Object indexes are automatically learnt from examples. As we have already commented, six images are acquired for each product, so that we have 6 (or 12 in the case of square boxes) invariant representations for each object. We then create a set of three indexes for each representation corresponding to three kinds of image features: color, appearance and edges. Indexes are computed using principal component analysis.

Principal component analysis ([Jolliffe 1986]) is a dimension reduction method which its first goal is to minimize the dimension of n-dimensional vectors to m-dimensional vectors (where $m < n$). PCA can be seen as a linear transformation that extracts a lower dimensional space that preserves the major linear correlations in the data and discards the minor ones. Vector projections can be used as representatives of original vectors for recognition purposes.

Having a data vector $\vec{x}$ ($\vec{x} \in R^n$), PCA projects it onto the m ($m < n$) dimensional linear subspace spanned by the leading eigenvectors of the data covariance matrix:

$$\Sigma = E[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^T],$$

where $\vec{\mu} = E[\vec{x}]$ and E denotes an expectation with respect to $\vec{x}$. The leading m eigenvectors $\{e_i | i_1, \ldots, i_m\}$ of a positive semidefinite matrix are the m eigenvectors which correspond to the m largest eigenvalues. The indices are assigned such that the corresponding eigenvalues in decreasing order are given by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$.

Let be $f : R^n \rightarrow R^m$ an encoding function from a vector $\vec{x} \in R^n$ to a vector $\vec{z} = f(\vec{x}) \in R^m$, where $m < n$. Let be $g : R^m \rightarrow R^n$ a decoding function from $\vec{z}$ to $\vec{x}' = g(f(\vec{x})) \in R^n$, a reconstruction of $\vec{x}$. PCA encodes $\vec{x}$ as:

$$\vec{z} = f(\vec{x}) = V(\vec{x} - \vec{\mu}) = (e_1^T(\vec{x} - \vec{\mu}), \ldots, e_n^T(\vec{x} - \vec{\mu}))$$

where V is an $m \times n$ matrix whose rows, e_i are the leading m orthonormal eigenvectors of Σ and $\vec{z}$ is the m dimensional encoding. The components of $\vec{z}$ are called the principal components. PCA reconstructs/decodes $\vec{x}'$ from $\vec{z}$ as:

$$\vec{x}' = g(\vec{z}) = V^T\vec{z} + \vec{\mu}$$

The mean squared error in reconstructing the original data is:

$$\epsilon = ||\vec{x} - g(f(\vec{x}))||^2$$

Figure 5: Object segmentation. (a) Original image, (b) Segmented blobs, (c) Hough transform of blob edges, (d) Detected rectangle, (e) Normalized image, (f) Sampled (32 × 32) normalized image, (g) Invariant image representation.

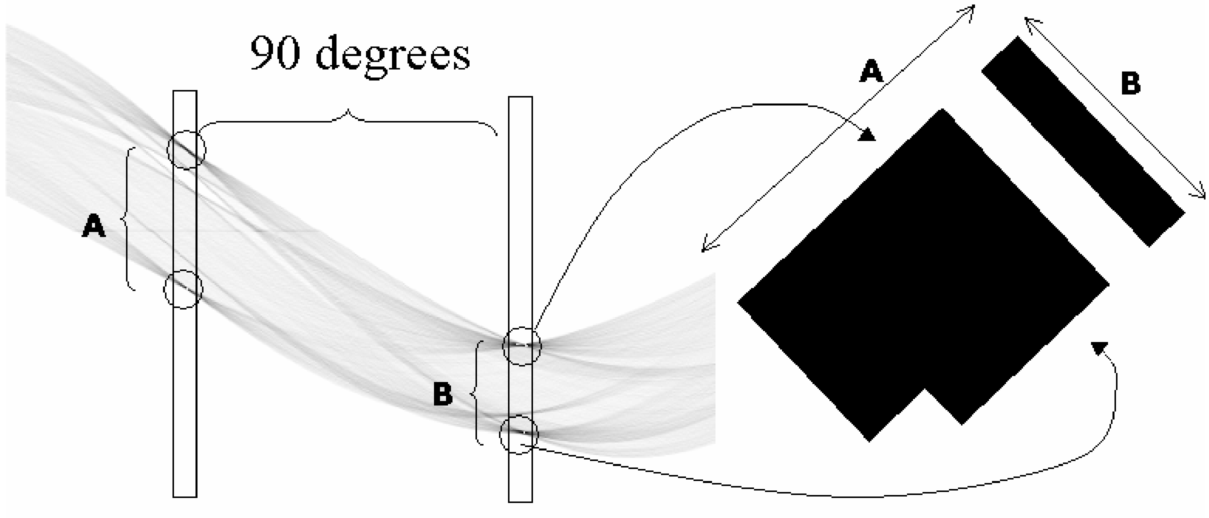

Figure 6: Representation of the Hough transform a of a segmented box. In this image the relation between the desired measures and the maxima of the accumulation space can be easily seen.

Using PCA we obtain the least error in terms of reconstructing the original vector.

PCA involves the computation of eigenvalues and eigenvectors of the data covariance matrix. When dealing with large high dimensional data sets, this computation can be difficult due to space requirements and complexity ([Oja 1983]). Due to this fact, we have considered an alternative representation, based on splitting the data in several groups and computing an encoding function for each group. This strategy has two advantages: the learning step can be performed more efficiently, and encoding functions are more adapted to objects, since they are computed for classes with fewer members.

In our case, the natural way of splitting data is based on size clustering. Boxes are distributed in size space in a non uniform way, describing different clusters (see figure 7a). In order to obtain a hard partition of this space, we have used a version of the watershed algorithm with markers ([Vincent and Soille 1991]). Markers have been selected as the 60 local maxima of the image that have the higher dynamics ([Grimaud 1992]). This characteristic, defined in the framework of mathematical morphology, is a good alternative to classical hard partition methods such as k-means algorithm.

Using this method we get 60 different clusters, from which we compute 60 different groups of descriptors. In the following section we describe the specific treatment for each descriptor in order to get the indexes. See Appendix I for a description of the clustering process.

3 Object representation and recognition

Each one of the 6 views of the objects in the database is stored as an array $(p_i, d_i, \vec{c}, \vec{a}, \vec{e})$, containing the following information: p_i product indentifier, d_i cluster identifier, $\vec{c}$ color index, $\vec{a}$ appearance index and $\vec{e}$ edges index. The cluster is determined by object size, and feature indexes are computed using the encoding functions learned from the set of images corresponding to that cluster.

3.1 EigenHistograms: statistical modeling of color distributions

Colors in an image are mapped into a discrete color space containing n colors. A color histogram is a vector in a n-dimensional space where each element represents the number of pixels of color j in the image. Given m images corresponding to different objects, we compute the set of histograms that

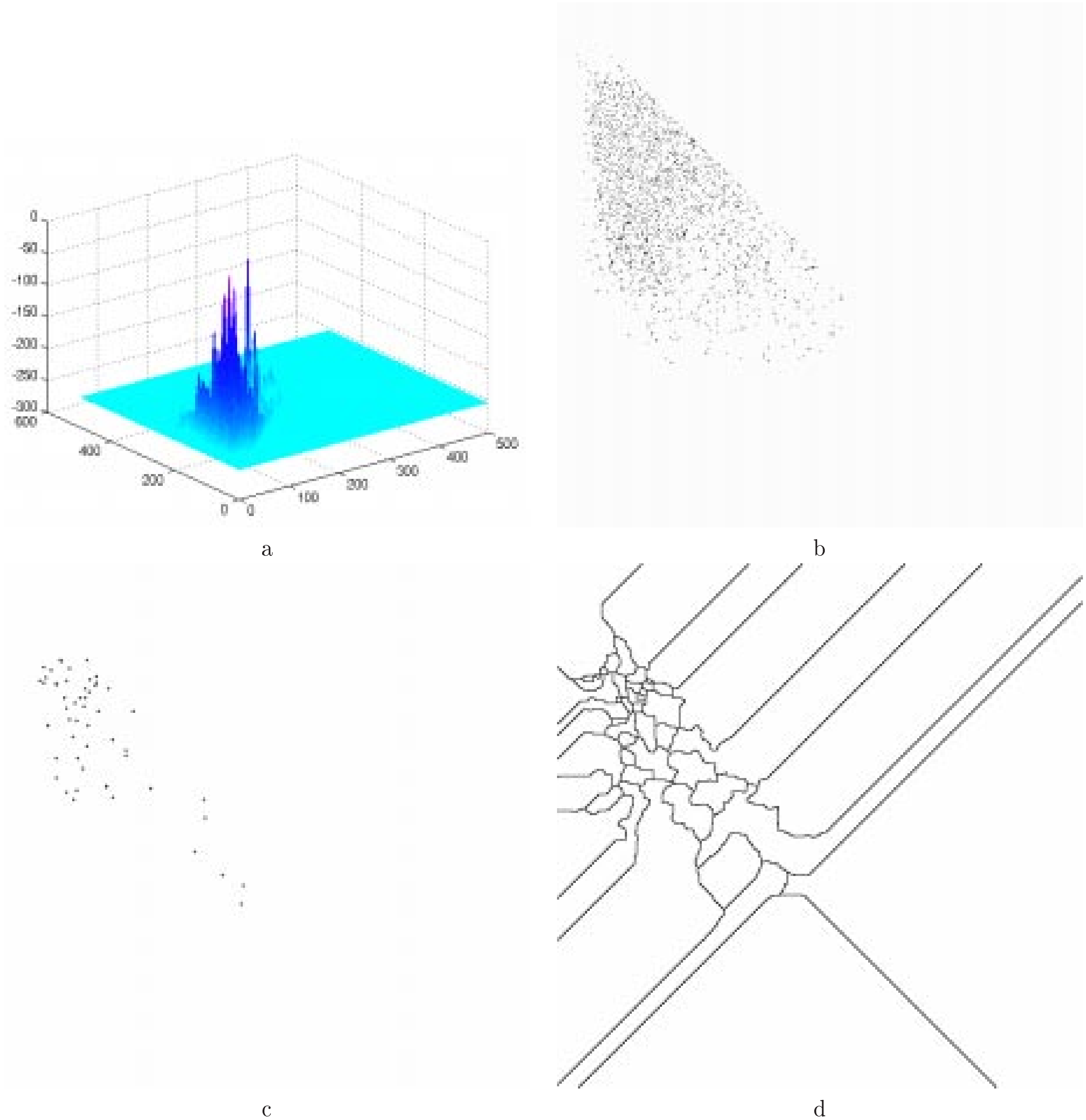

Figure 7: Object clustering. (a) Size distribution, (b) Extrema of the size distribution, (c) Extrema with the highest dynamics and (d) Watershed segmentation of the size space.

represents them as follows:

$$O = \{H_1, \ldots, H_m\}$$

This set of histograms, the learning set, represents all possible variations of the objects color distributions. We assume that these variations span a low dimensional linear subspace of R^n (being n the number of colors represented in the histogram) which can be computed using PCA ([Vitrià et al. 1999]). Principal component analysis of a histogram set defines an encoding function that projects each histogram from the set onto a low dimensional subspace defined by the k, being $k < m$, principal eigenvectors e_i of the histogram covariance matrix. Given that $\{e_i | i_1, \ldots, i_m\}$ forms an orthonormal base for object histograms, we call each e_i an *EigenHistogram*.

To identify an object from its color distribution, the k significant EigenHistograms are chosen as those with the largest associated eigenvalues. Then, the histogram corresponding to color distribution is transformed into its EigenHistogram components, c_i. These components form a vector $\vec{c}$ that describes the contribution of each EigenHistogram in representing the input color distribution. The component vector is classified as belonging to an object j when $\vec{c}$ minimizes the Euclidean distance

$$\epsilon_j = ||\vec{c} - \vec{c}_j||^2$$

where $\vec{c}_j$ is the component vector resulting from projecting a histogram of the learning set that corresponds to object j.

3.2 Appearance and edges

In the absence of color information or in the presence of images with similar colors it becomes necessary the use of additional image features for product recognition. We describe objects by analyzing two related characteristics: significant edges and appearance.

The 32 by 32 normalized image of a pharmaceutical product is considered as a vector of dimension 1024. The set of normalized images corresponding to the products of a cluster maps to a collection of points in this space. In order to recognize a product from its appearance, the projection $\vec{a}$ of the normalized image vector to the low dimensional subspace formed by the eigenvectors of the set of training images can be used. As in the case of color histograms, objects can be recognized by minimizing the Euclidean distance between $\vec{a}$ and the component vectors of the cluster learning set members.

Significant edges of the product image can also be used in this framework to increase recognition capabilities of the system. Considering the class of objects we are dealing with, the use of color and appearance may not be enough to differentiate between similar products: the presence of a small detail on the surface of a product box can be determinant to correctly identify it. Hence, we have added a descriptor sensible to this information. The method consists of the following steps:

1. Edges are detected in the 32×32 normalized image using the Roberts operator.

2. The invariant image representation is created.

3. Vertical and horizontal projections are taken, resulting in two histograms, that are concatenated in a unique descriptor.

4. Principal Component Analysis of the set of edge descriptors corresponding to all the images of a cluster defines a encoding function for this descriptor, that results in a component vector $\vec{e}$.

5. As in the previous cases, objects can be recognized from edge projections by comparing its component vector to the cluster learning set.

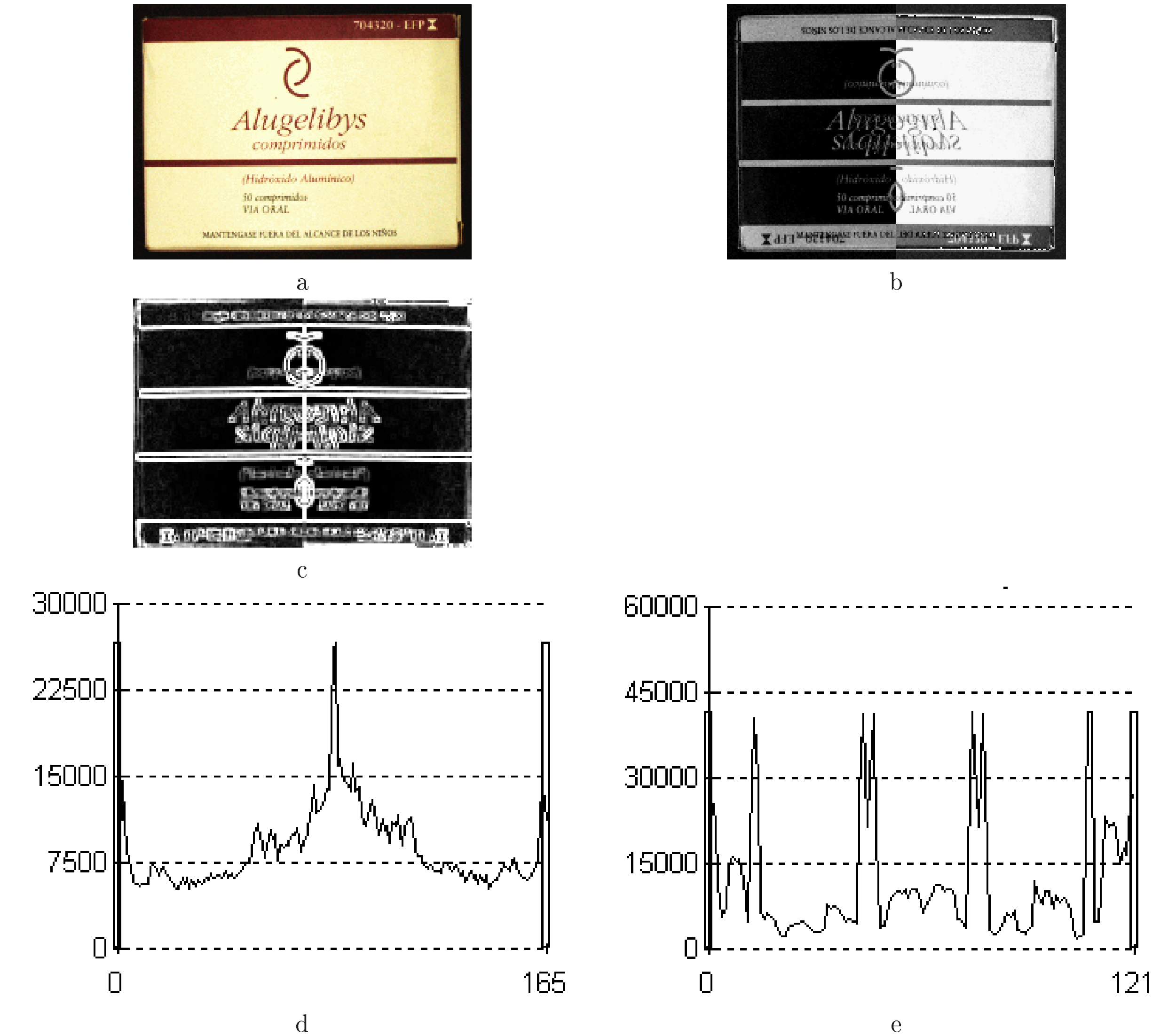

Figure 8: Edge information: (a) Product image, (b) Invarint image representation, (c) Edges on the invariant image representation, (d) Vertical projection, (e) Horizontal projection.

Table 1: Retrieval results (%) using color, appearance, edges, and their integration. n refers to the position of the correct product.

Query nature	n=1	n=2	n=3	n=4	n=5
Color	86.05	93.82	96.16	97.30	97.9
Appearance	89.70	95.67	97.42	98.25	98.8
Edges	89.34	95.73	97.56	98.36	98.78
Integrated query	92.86	96.91	98.03	98.64	99.05

3.3 Product recognition

When an unknown object x is presented to the system, its identification process proceeds as follows:

1. The image is segmented, and using the modified Hough transform, the size and localization of the object are computed. Object cluster is determined from size.
2. The segmented image is normalized.
3. The object descriptor $(\vec{c}_x, \vec{a}_x, \vec{e}_x)$ is computed using the eigenvectors of each considered feature (color, appearance and edge) corresponding to its cluster.
4. The object descriptor is compared, using an Euclidean distance, to all the product descriptors which belong to its cluster. A three-nearest neighbor rule is applied in order to decide which is the most probable product.

Special attention must be made to products that have a box size corresponding to a point in the size space that is near a cluster frontier. In this case, and due to variations in object position in the field of view, the measure of object size can suffer of small errors, and a wrong cluster can be selected. In order to avoid this problem, we have defined an overlapping region between clusters. Each cluster has a central region where confidence is high, and several regions where misclassification is possible due to the proximity of another cluster. Objects that belong to these regions [2] are considered to belong to all the clusters that share them. Hence, instead of representing the object by one object descriptor, the object is represented by n object descriptors, being n the number of clusters that are near its point in the size space. The minimization of the distance between one of these descriptors and the set of models determines the object identity.

This recognition scheme allows for the processing of objects in less than a second.

4 Experimental evaluation

We have extensively tested the accuracy and stability of our image indexing system. Experiments were conducted on the entire database of 7458 products (49056 reference images). Table 1 presents the results of retrieval based on color, appearance and edges, respectively, where n refers to the position of the correct product retrieval.

We designed the test in this manner: we selected 5 images of each product for computing optimal feature subspaces, and 5 descriptors were stored in the model database for each product. The sixth image was presented as a query. This process was repeated for all combinations of learning set and query image. Results in Table I shows the results averaged over all the tests.

[2] Note that by construction, frontier regions are the less populated regions of the size space.

Retrieval performance of appearance is slighly better than color or edges when considered independly, although is comparable. When all features are considered, we get an increasing of retrieval performace. The number of false matches for the first retrieval is due to the presence of very similar products in the database, but this problem is solved when considering the top five retrievals.

5 Conclusions

We have presented a prototype system for indexing pharmaceutical products. We have defined a descriptor for each product that is based on learning optimal feature subspaces from examples. Selected features included color, appearance and edges. Size clustering was used to increase system discriminative performance, and to decrease computational requirements. The system allows for high retrieval efficiency in databases with thousands of different objects. Future works deals with comparing our approach to alternative subspace learning methods for feature representation (such as Independent Component Analysis), and studing different metrics for recognition.

A Mathematical morphology based clustering

Consider a function $f : R^n \Rightarrow R$. We define the local maxima of f as a connected set of points $\{x_1, ..., x_n\}_{n \geq 1}$ such that $f(x_1) = f(x_2) = ... = f(x_n)$ and all their neighbors have lower values. In a clustering process, we can associate the maxima of f with the center of the classes, and the influence zones of each maximum with the support set of each class. The influence zones are defined as zones of the feature space where we can reach the maximum by following an increasing path starting from any point. These zones can be determined by applying the watershed detection algorithm ([Vincent and Soille 1991]) to the complement of the function f.

To avoid an oversegmentation of the feature space, a variation of the watershed algorithm can be used. In this case, significant maxima of the feature space are selected by an additional criterion. The modified watershed algorithm, known as *watershed detection with markers*, modifies f in order to get a new function f' where the only local maxima are those selected by the marking process, but the topography of f has been maximally preserved.

In order to select the maxima of interest, we have used a transformation that valuates the extrema of the function on a contrast criterion: the dynamics (Grimaud, 1992). The dynamics of a maximum can be defined using notions such as watersheds and influence zones.

The following definitions allow for a clear definition of the dynamics of a maximum:

Definition A.1 *A* path $P(x, y)$ *on* f *is a set of neighbor points* $(x_0, x_1, ..., x_n)$ *that links point* x *to* y*.*

Definition A.2 *The* dynamics of a path *points* x *and* y *on* f *is the absolute difference between* $f(x_h)$ *and* $f(x_l)$*, where* x_h *is a point such that* $f(x_h) \geq f(x_i)$*, and* x_l *is a point such that* $f(x_l) \leq f(x_i)$ *for every point* x_i *in the path.*

Definition A.3 *The* dynamics between two points x *and* y *is equal to the dynamics of the path with the lowest dynamics between* x *and* y*.*

Definition A.4 *The* dynamics of a maximum M *is equal to the dynamics of the path with the lowest dynamics that links the maximum* M *to a pixel* y *such that* y *belongs to a influence zone whose maximum has a strictly lower value than* M*.*

Extrema dynamics and watershed detection can be efficiently computed using a technique called *immersion simulation* ([Vincent and Soille 1991]).

Acknowledgement 1 *This work was partially supported by CICYT and EU grants TAP98-0631, TIC98-1100, TAP97-463 and 2FD97-0220, and by CERF Catalunya S.A.*

References

[Black and Jepson (1996)] Black M., Jepson A. (1996). Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. Proceedings of ECCV, pp.329-342.

[Buhman et al. (1990)] Buhmann J., Lades M., von der Marlsburg C.(1990) Size and distortion invariant object recognition by hierarchical graph matching. Proc. IEEE IJCNN, San Diego, pp. 411-416.

[Funt and Finlayson 1995] Funt B., Finlayson G. (1995) Color Constant Color Indexing. IEEE Trans. PAMI, 17, 5, pp.522-528.

[Gevers and Smeulders 1997] Gevers T., Smeulders A. (1997) Color Based Object Recognition. In Image Analysis and Processing, Alberto del Bimbo (Ed), LNCS 1310, pp. 319-326.

[Grimaud 1992] M.Grimaud M. (1992) A new measure of contrast: the dynamics. SPIE Vol. 1769, Image Algebra and Morphological Image Processing III, pp. 292-305.

[Grimson 1990] Grimson W. (1990) Object Recognition by Computer, MIT Press.

[Huttenlocher and Ullman 1987] Huttenlocher D.P., Ullman S. (1987) Recognizing Solid Objets by Alignment. Proc. IEEE International Conference on Computer Vision, pp 102-111.

[Jolliffe 1986] Jolliffe I.T. (1986) Principal Component Analysis. Springer Verlag, New York.

[Lowe 1987] Lowe D.G. (1987) Three-Dimensional Object Recognition From Single Two Dimensional Images. Artificial Intelligence, 31:355-395.

[Martinez and Vitrià 1997] Martinez A., Vitrià J. (1997). Dimensionality Reduction for Face Recognition, in Advances in Visual Form Analysis, pp. 405-414, World Scientific, Singapore.

[Moghaddam and Pentland 1996] Moghaddam B. and Pentland A. (1996) Probabilistic Visual Learning for Object Recognition, in Nayar S. and Poggio T. (eds.) ”Early Visual Learning”, Oxford University Press.

[Murase and Nayar (1993)] Murase H., Nayar S.K. (1993) Learning and recognition of 3D objects from appearance. Proc. IEEE Qualitative Vision Workshop, New York, pp. 39-49.

[Oja 1983] Oja E. (1983) Subspace methods of pattern recognition. Res. Studies Press, Hertfordshire.

[Phillips et al. 1998] Phillips P.J., Moon H., Rizvi S., Rauss P. (1998) The FERET evaluation. In Face Recognition, from thery to applications, Springer, berlin, 1998.

[Rao and Ballard 1995] Rao R., Ballard, D. (1995) Object Indexing using an Iconic Sparse Distributed Memory. In Proc. of the ICCV, pp.24-31.

[Schielle and Crowley 1996] Schielle B., Crowley J.L. (1996). Object recognition using multidimensional receptive fields histograms. In Proc. of ECCV, pp.610-619.

[Swain and Ballard 1991] Swain M., Ballard D. (1991) Color Indexing. Intern. J. of Computer Vision, 7, 1, pp. 11-32.

[Turk and Pentland (1991)] Turk M.A. and Pentland A. (1991).Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3 (1), 71-86.

[Ullman 1996] Ullman S. (1996) High-Level Vision, MIT Press.

[Vincent and Soille 1991] Vincent L. and Soille S. (1991) Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13,no.6,pp.583-598.

[Vitrià et al. 1999] Vitrià J., Radeva P. and Binefa X. (1999) EigenHistograms: using low dimensional models of color distribution for real time object recognition. 8th International Conference on Computer Analysis of Images and Patterns, Springer Verlag LNCS 1689, pp. 17-24.

Biography 1 *Jordi Vitrià joined the Computer Science Department of the Universitat Autònoma de Barcelona (UAB) in 1986, becoming associate professor in 1991. In 1990 he received the Ph.D. at the UAB on work in mathematical morphology and image analysis. He has collaborated in the creation of the Computer Vision Center (CVC), a R&D institute founded by the UAB and the Generalitat de Catalunya (autonomous government of Catalonia). J.Vitrià is an active member of the Image Analysis and Pattern Recognition Spanish Association (AERFAI), a branch of the IAPR, where he has coordinated several activities. He is cofounder and former Secretary of the Catalan Association for Artificial Intelligence (ACIA) (associated to the European Coordinating Committee on Artificial Intelligence -ECCAI-), an organization created in 1994 to promote and coordinate AI activities in Catalonia. He has been since 1991 the organizer of several activities related to image analysis, computer vision and artificial intelligence. His main research interests include object recognition and statistical techniques applied to computer vision.*

Biography 2 *Petia Radeva has graduated in Faculty of Mathematics and Computer Science, University of Sofia, Bulgaria in 1989 and received her M.Sc. and Ph.D. degrees in Computer Science at the Universitat Autonòma de Barcelona, in 1993 and 1996. She has been coordinating a national research project on Computer Vision applied to Medicine as well as private projects on industrial applications of image processing tools. In 1997-1998 she has been a national contact point of Spain of a concerted action "MAVIRIC" sponsored by the European Union. Her particular interests involve 3D reconstruction of flexible models and recovering of non-rigid motion, on which topics she has published several international papers.*

Biography 3 *Xavier Binefa received de Ph.D. degree in Computer Science from the Universitat Autònoma de Barcelona in 1996. He is currently a lecturer in the Department of Computer Science of that university. His present research interests include computer vision and multimedia technologies.*

Biography 4 *Albert Pujol received the M.Sc degree in Computer Science from the Universitat Autònoma de Barcelona in 1998. He is currently an assistant professor in the Department of Computer Science at the Universitat Autònoma de Barcelona. His research work has been related to invariant object recognition and neural networks, and he is actually working in his Ph.D. Thesis on view point invariant action recognition.*

Biography 5 *Ernest Valveny received the B.Sc. degree in computer science from the Universitat Politècnica de Catalunya in 1992, and the M.Sc. degree from Universitat Autònoma de Barcelona in 1994. He is an assistant professor in the Computer Science Department at Universitat Autònoma de Barcelona since 1992, and member of the Computer Vision Center. His research interests are in the field of hand-drawn document interpretation and analysis using techniques based on deformable template matching.*

Biography 6 *Robert Benavente received the B.Sc. degree in Computer Science from the Universitat Autònoma de Barcelona in 1998. Presently, he is an assistant professor in the Computer Science Department at the Universitat Autònoma de Barcelona and a research member of the Centre de Visio per Computador (CVC). His research interests, in order to obtain his M.Sc. degree, include Colour Perception and Texture Discrimination. In addition, he participates in the development of machine vision applications for the industry.*

Biography 7 *Craig von Land received his B.Sc. degree in Physics from the University of Alberta, Canada in 1986, and completed his M.Sc. degree in Computer Vision at the Universitat Autònoma de Barcelona, Spain in 1998. He has led the development of several commercial medical imaging products, and is currently working as a senior software design engineer in medical imaging for GE Marquette Medical Systems in Jupiter, Florida.*