Bayesian Decision versus Voting for Image Retrieval (extended version accepted to CAIP 1997)

R Mohr, S. Picard, C Schmid

June 6, 1997

Abstract

Image retrieval from image databases is usually performed by using global image characteristics such as texture or colour. The use of local image information is highly desirable when only part of the image is of interest, but global approaches are not well suited to this. An original solution was introduced in [11] using invariant local signal characteristics. This paper extends this contribution by extending the set of invariants considered to allow illumination change. Then it is shown that the invariant distribution is far from uniform and a probabilistic indexing scheme is proposed. Experimental results validate the approch and the different method are discussed. The main result is that it is much more valuable to increase the discrimant power of the vector used to perform the indexing process; The Bayesian decision improves the standard method, but this improvement is much more limited than expected.

1 Introduction

In this paper we address the problem of retreiving images from a large database based on local measures of similarity. Similarity may be based on colour, texture, shape or gray level properties of the image under query as is illustrated for a pair of matching aerial images shown in figure 5. However, variations in viewpoint, scale or illumination complicate the problem significantly by requiring the definition of similarity measures which are invariant under such changes. In this paper we extend the use of local gray level invariant features into a probabilistic framework.

While in many applications, one would like to allow database query requests like "*find the van Gogh painting with a dancing person*", the state of the art in image processing is far from what is needed in order to allow such high level requests. Therefore existing query systems are usually interactive, allowing the user to search for similar images based on a chosen image of interest. Scientists might consider that such systems are poor in what they offer, but such tools are already useful for intelligent image data base browsing. Database query on the basis of color histograms [14], of texture [5], greylevel processed information (gradient, etc) [9] or Principal Component Analysis [15] have all achieved some success and some commercial products have been developed based on these ideas (see web sites for Virage, Qbic or Excalibur for instance).

However, the above mentioned techniques are all based on the global analysis and comparison of images and therefore are not well suited to matching when images of an object are taken under different viewing conditions. We are interested in the development of image database query techniques which can address the problem of locating a pattern in a subimage of a database image when that pattern is obtained from different viewing conditions:

- partial visibility
- different viewing angles
- complex scene with complex background
- many potential reference pattern.

Under such conditions, data have to extracted locally, and the general process will have to cope with missing and spurious data.

In this paper we begin by describing the method by which locally invariant characteristics can be extracted from image (section 2.1), and how an indexing scheme based on these features can provide fast access to the relevant corresponding images. This section extends mainly the work presented by Schmid in [11] by considering not only invariance in geometry but also in illumination. The next section provides the Bayesian analysis of the retrieval process and, based on an estimated density of the characteristic in the image, provides a Bayesian scheme for image retrieval. This leads to experiements which compare the new scheme to the basic scheme and its extension with semi local characteristics presented in [11].

2 Indexing with greylevel invariants

The goal of indexing is to locate all instances of images I of a scene in a database. However, the actual images I will be different due to variations in viewing conditions:

1. change of the position of the camera observing the scene
2. camera with different gain functions
3. change of illumination
4. occlusion and clutter

Therefore, it is imperative that the properties used as the basis of any indexing scheme be invariant with respect to these changes. Although it has been pointed out by several authors that there are no generally invariant features, even for a change in the observer position (see for instance[2]), our goal is to find realistic simplified assumptions for computing stable signal properties.

2.1 Geometric greylevel invariants

For shape oriented towards the observer, similarities (i.e. rotation, translation and scaling) absorb the first order of geometric changes with respect to camera motion [1]. Therefore computing invariants for this group of image displacements would address the problem of change in camera position. However, in order to be robust in the presence of occlusion and clutter the computation of global surface invariants [13] must be abandoned in favour of locally invariant features.

As explained in [11] we employ differential invariants that have been theorically studied by Kœnderink [4], Romeny *et al* [8, 3, 6, 16]. Interested readers are referred to these works. These greylevel invariants are computed at points of interest, typically corners, ensuring that they are located in regions where the signal is informative.

Following [16] we compute the complete set of differential invariants up to the third order. The set of invariants is stacked in a vector, denoted by $\vec{X}$ which is broken into two parts. The first part of this vector contains the complete and irreducible set of differential invariants up to 2nd order (cf. equation

1). The formulation of this first part of the vector is given in so-called Einstein notation.

$$\vec{X}[0..4] = \begin{bmatrix} L \\ L_i L_i \\ L_i L_{ij} L_j \\ L_{ii} \\ L_{ij} L_{ji} \end{bmatrix} = \begin{bmatrix} L \\ \frac{\partial L^2}{\partial x} + \frac{\partial L^2}{\partial y} \\ \frac{\partial L^2}{\partial x}\frac{\partial^2 L}{\partial x^2} + 2\frac{\partial L}{\partial x}\frac{\partial L}{\partial y}\frac{\partial^2 L}{\partial x \partial y} + \frac{\partial L^2}{\partial y}\frac{\partial^2 L}{\partial y^2} \\ \frac{\partial^2 L}{\partial x^2} + \frac{\partial^2 L}{\partial y^2} \\ \frac{\partial^2 L}{\partial x^2}^2 + 2\frac{\partial^2 L}{\partial x \partial y}\frac{\partial^2 L}{\partial y \partial x} + \frac{\partial^2 L}{\partial y^2}^2 \end{bmatrix} \tag{1}$$

L represents a gaussian kernel convolved with the image and the L_i are the first derivatives of the luminance according to i, which represents either x or y coordinates. The derivatives are computed by convolving the image with derivatives of the gaussian kernel. This vector can be considered as the projection of the image information onto a set of basis functions.

The second part of the vector contains a complete set of invariants of third order. Equation 2 presents this set limited to the Einstein notation.

$$\vec{X}[5..8] = \begin{bmatrix} \varepsilon_{ij}(L_{jkl}L_i L_k L_l - L_{jkk}L_i L_l L_l) \\ L_{iij}L_j L_k L_k - L_{ijk}L_i L_j L_k \\ -\varepsilon_{ij}L_{jkl}L_i L_k L_l \\ L_{ijk}L_i L_j L_k \end{bmatrix} \tag{2}$$

with ε_{ij} the 2D antisymmetric Epsilon tensor defined by $\varepsilon_{12} = -\varepsilon_{21} = 1$ and $\varepsilon_{11} = \varepsilon_{22} = 0$.

At this stage we therefore have, for each point of interest, a vector $\vec{X}$ of 9 values which are invariant to rotation and translation.

Invariance in scale can be computed too. This leads however to change also the local support for which the computation is done in the image. Although in theory such invariants exist, in practice they have to be computed at different scales and stacked in the indexing structure.

2.2 Illumination invariance

Illumination changes, both in direction and in intensity produce strong changes in an image, particularly where shadows are moving. Although there are no invariants which can be computed under such general changes in illumination, we *can* consider the simplified case in which we ignore the problems associated with shadows by considering that the surface around the point of interest is locally flat and that its brightness (irradiance) is computed as the product of the local surface reflectance function $\rho(u,v)$ with some factor k. k depends on the intensity and direction of the light source and on the normal orientation of the surface. Under our assumption k is constant around a point of interest in a specific image, thus capturing the usual reflectance model, which includes the assumption of Lambertian reflectance.

For an ideal camera, this model implies that changes in illumination would cause the pixel intensity to be scaled by an unknown local factor. However, real camera transfer functions look more like the one displayed in Fig 1, meaning that a linear assumption can be made almost out of the extremal values. The resulting intensity displayed in each pixel will not be an affine version of the real one.

$$\begin{aligned} I_{\text{image}} &= \alpha I_{\text{input}} + \beta \\ &= \alpha' \rho_{\text{object}} + \beta' \end{aligned}$$

The two factors $\alpha'\beta'$ have to be cancelled out of our measures. The offset β' is eliminated from all of the measures except the first one since they are based on spatial derivatives. The uniform scaling β' is eliminated by considering ratio's of the remaining differential invariants. Illumination invariance is therefore achieved by replacing the vector of differential invariants $\vec{X}[0..8]$ with $\vec{Y}[2..8]$:

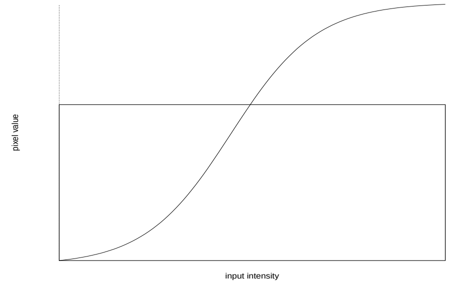

Figure 1: Typical camera transfer function

$$
\begin{aligned}
\vec{Y}[2] &= \frac{\vec{X}[2]^{\frac{2}{3}}}{\vec{X}[1]}; & \vec{Y}[3] &= \frac{\vec{X}[3]}{\vec{X}[1]}; & \vec{Y}[4] &= \frac{\vec{X}[4]}{\vec{X}[1]}; \\
\vec{Y}[5] &= \frac{\vec{X}[5]}{\vec{X}[1]^2}; & \vec{Y}[6] &= \frac{\vec{X}[6]}{\vec{X}[1]^2}; & \vec{Y}[7] &= \frac{\vec{X}[7]}{\vec{X}[1]^2}; \\
\vec{Y}[8] &= \frac{\vec{X}[8]}{\vec{X}[1]^2};
\end{aligned} \tag{3}
$$

Images can be matched under very different illumination conditions using these invariants. Fig 3 illustrates the matching performance under illumination changes using using $\vec{X}$, just when canceling the β and when using the illumination invariant features $\vec{Y}$. Illumination varies widely as displayed in Fig 2.2.

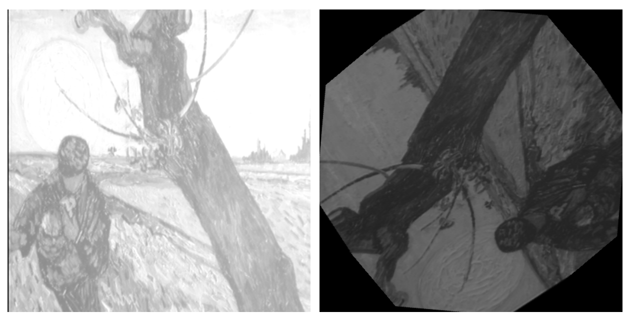

Figure 2: Example of a change of illumination

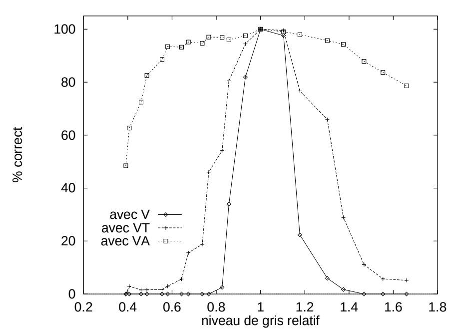

Figure 3: Successful matches when using $\vec{X}$ (see V), when canceling the β (VT) and using $\vec{Y}$ (VA).

2.3 A Basic indexing algorithm

Local invariant vectors contain provide useful information that can be used for indexing image objects in a manner similar to the way geometric characteristics have previously been used [7]). However, the local points of interest from which these vectors are computed cannot be robustly detected in the presence of significant changes in the observed images. Typically only about 50% of the points of interest will be detected in any pair of images of an object. Therefore an index alone is meaningless; moreover, when a large number of such points occurs in an index table (we have more than 150,000 such vectors for our data base), ambiguity may arise. For this reason, Schmid in [12] used a voting technique similar to the Hough transform in which multiple feature matches are used to form a consensus of which database images provides the best match to a query.

The basic algorithm is therefore:

- compute the vector of invariants at each point of interest;
- compare them to invariants stored in the database (matching) using a fast access data structure and record the number of matches found for each database image;
- retrieve the image with largest number of matches.

This method was successful for retrieving images taken under new conditions with significant changes.

3 A probabilistic model for the image retrieval

3.1 Invariant distribution

The basic indexing algorithm described above assumes that all vectors are equally probable, therefore according equal importance to all matches. However, this is not necessarily the case. Following a suggestion made by Schiele [10], we observed the distribution of the invariants collected in our image data base, which contains over 150,000 9-dimensional vectors. Figure 4 represents the repartition of invariants 5 and 6, showing the nature of the distribution. It is clear the the information associated with each point is highly different and should be taken into account.

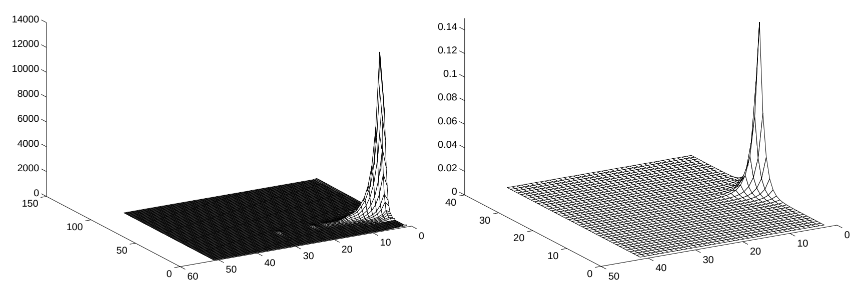

Figure 4: Distribution of the invariants 5 and 6 in the data base (left) and estimated (right)

Given that the observed distribution is unimodal, we model it as a product of unimodal 1-dimensional density functions. The non negative invariants were modelled with a Weibull distribution, all other invariants were modelled using Gaussian or Laplacian distributions. It has been verified that these distributions fit the invariant statistics. The resulting distribution almost perfectly matches the input statistics, as shown in figure 4b.

3.2 The basic Bayesian model

In a study on the distribution of the measure in some “receptive fields”, Schiele [10] derived a Bayesian model which allowed the distribution of the measures to be taken into account. In this section we take a similar approach, deriving a model which uses the a priori knowledge of the distribution of the invariants in the matching process. But, as this work is concerned with matches or potential matches, we have to derive a more sophisticated model in order to take into account the matching process.

Let Q be the query image and R be a database image being considered as a match candidate. It is assumed that Q and R will have a large number of features in common if they are a correct match. Q has $n(Q)$ interest points, and R has $n(R)$ such points. From the $n(Q)$ features of Q, $\{m_i\}, i \in I$ is the set of the features which are matched with features in R. Each m_i is a feature vector of Q that has a matching feature vector f_k appearing in R.

We want to evaluate $P(R|\{m_i\})$. Using Bayes formula we get

$$P(R|\{m_i\}) = \frac{P(\{m_i\}|R).P(R)}{P(\{m_i\})} \tag{4}$$

Assuming that the individual matches are independent, this translates to

$$P(R|\{m_i\}) = \frac{\prod_{i\in I} P(m_i|R)P(R)}{\prod_{i\in I} P(m_i)} \tag{5}$$

$P(m_i)$ is the probability that the i-th feature of Q has one match with $n(R)$ random features.

This approach considers only the effect of matches, but the fact that many features fail to match must also be considered. To incorporate this let $\{\overline{m}_j\}, j \in J$ be the set of of features that failed to be matched. Formula (5) can thus be extended:

$$P(R|\{m_i\},\{\overline{m}_j\}) = \frac{P(\{m_i\},\{\overline{m}_j\}|R).P(R)}{P(\{m_i\},\{\overline{m}_j\})} \tag{6}$$

which becomes

$$P(R|\{m_i\},\{\overline{m}_j\}) = \frac{\prod_{i\in I} P(m_i|R)\prod_{j\in J} P(\overline{m}_j|R)P(R)}{\prod_{i\in I} P(m_i)\prod_{j\in J} P(\overline{m}_j)} \tag{7}$$

3.3 Probabilistic interpretation of direct voting

Counting the number of feature matches between a query and each database images and selecting the image(s) R for which the number of matches is maximum is an effective strategy for matching. Impressive results have been obtained using this strategy, explained by the fact that this counting approach is robust to outlier matches [11]. In this section we consider this matching in a a probabilistic framework.

Let us consider each feature in Q as the result of a uniform random variable whose values can be one of the values observed in all the possible Rs. The probability p_i^R that the i-th feature of Q matches one of the feature knowing that Q is part of R is the probability of repeatability of the feature in R when viewing conditions varies. Such a probability can be considered as a constant α; experimentally, depending of the change of illumination and of viewing conditions, α has been estimated to be in the range [0.2..0.5].

Similarly, the probability $P(m_i)$ will be some constant β. $\beta \ll \alpha$ as the number of features possible in all the data base is much larger than the one appearing in R. This leads to :

$$P(R|\{m_i\},\{\overline{m}_j\}) = \frac{\prod_{i\in I}\alpha\prod_{j\in J}(1-\alpha)}{\prod_{i\in I}\beta\prod_{j\in J}(1-\beta)}P(R) \tag{8}$$

As the logarithm is a monotic function, the largest probability over R is the largest logarithm.

$$\begin{aligned}\log(P(R|\{m_i\},\{\overline{m}_j\})) &= \sum_{i\in I}(\log\alpha - \log\beta) + \sum_{j\in J}(\log(1-\alpha) - \log(1-\beta)) + \log(P(R)) \\ &= \mid I \mid (\alpha' - \beta') + \mid J \mid (\gamma' - \delta') + \log(P(R))\end{aligned}$$

As $| I | + | J |= n(Q)$, $\log \alpha - \log \beta > 0$ whereas $\log(1-\alpha) - \log(1-\beta)) < 0$. Assuming all $P(R)$ equal, this clearly shows that the maximum is obtained when $| I |$ is maximum, i.e. the number of matches is maximum.

This means that assigning a match based on the maximum vote corresponds to a maximum likelihood decision under the assumption of a uniform distribution of features on the object and on matches.

3.4 Posterior Probability of Retrieved Images

Matching also occurs randomly, inducing false matches, and this was not taken into account in the previous model. The model introduced here will consider:

- correct matches (i.e. corresponding to the good image)
- false matches (i.e. corresponding to a random process)

If we assume that Q is a subimage of R under some new viewing condition, the k-th feature of Q might be a feature of R with the previously defined probability α. It could be also a feature that occured due to some random process with the density probability of R.

Let p_R^k be the probability of the k-th random feature vector of Q to appear in image R. p_R^k is estimated using the distribution described in 3.1, but just restricted to the particular image R. The probability that it might miss all the $n(R)$ features in R is therefore $(1-p_R^k)^{n(R)}$. Thus the corresponding probability of matching one of these features is $1-(1-p_R^i)^{n(R)}$. Combining this two events which are exclusive, the likelihood of match (a correct or a false one) becomes:

$$P(m_i|R) = \alpha + (1-\alpha)(1-(1-p_R^i)^{n(R)}) \tag{9}$$

Similarly the probability of occurence of the j-th feature vector of Q will be p_B^k (B for data *b*ase) and the a priori probability of match with $n(R)$ feature vector is then

$$P(m_j) = \beta + (1-\beta)(1-(1-p_B^i)^{n(R)})$$

with β defined in the previous subsection. Substituting this into equation 5 results in:

$$P(R|\{m_i\}) = \frac{\prod_{i\in I}(\alpha + (1-\alpha)(1-(1-p_R^i)^{n(R)}))}{\prod_{i\in I}(\beta + (1-\beta)(1-(1-p_B^i)^{n(R)}))} P(R) \tag{10}$$

The probability for the non matched point is modelled in a similar way leading to the refined version of equation 6:

$$P(R|\{m_i\},\{\overline{m}_j\}) = \frac{\prod_{i\in I}(\alpha + (1-\alpha)(1-(1-p_R^i)^{n(R)}) \prod_{j\in J}(1-\alpha)(1-p_R^j)^{n(R)}}{\prod_{i\in I}(\beta + (1-\beta)(1-(1-p_B^i)^{n(R)})) \prod_{j\in J}(1-\beta)(1-p_R^j)^{n(R)}} P(R) \tag{11}$$

3.5 Preliminary discussion

At this stage one may notice that if p_R^i is very low, then equation 9 simplifies to $P(m_i \mid R) = \alpha$, i.e. to the case when false matches are not considered. One might increase the local information by using much more information, leading to very low probability of false matches and to the simplified model defined by equation 8.

In [11] the authors used semi-local constraints for their voting technique: A point was considered as matched only if the invariant descriptor matched, and also if two additional matches were observed in the neighbourhood of the five closest points of interest. [11] reports that this semi local constraint was necessary for obtaining good results in the case of matching complex aerial images (see next section).

The goal of the next section will be to experimentally determine the utility of including the a priori knowledge of the data distribution in the matchin process.

4 Experiments

The experiments reported here are based on querying aerial images only (see fig. 5). As reported in [11], these are the most difficult images: They are similar in texture and shape, as roofs of old houses look very similar seen from the sky.

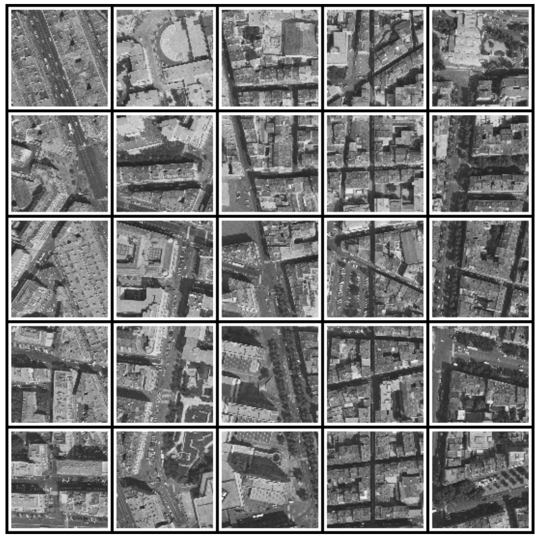

Figure 5: Some examples of the aerial image base

The query images were taken from an airplane from a position different from the reference images (about 20 degrees change in the almost vertical viewing direction). Altitude was the same, so the scaling effect can be neglegted. However due to the change of viewing direction, the images differ: some facades are visible, illumination has changed, etc. (see fig 6).

One hundred images corresponding to the first view point are stored in the database. Half of the hundred image of the second view point are used for the query.

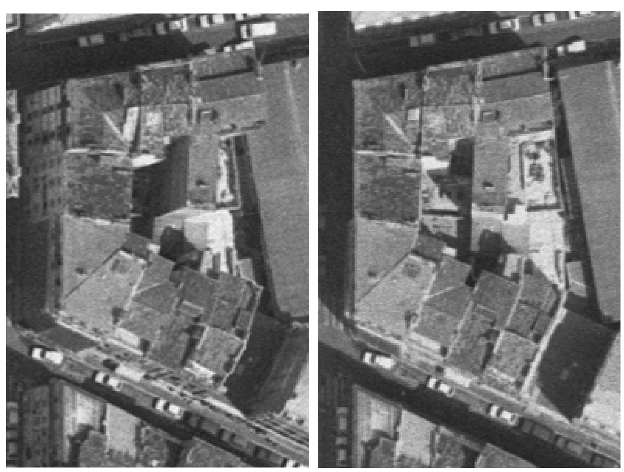

Figure 6: Part of an image of the base and the request image (query and model). Courtesy of ISTAR.

4.1 Experimental process

For each image, the probability p_R^i has been estimated just using a histogram, as an image does not provide enough data for estimating the probability density in the high dimensional space we are considering.

The value for α in the Bayesian model was set to 0.5 as experimentally this was close to the repeatability, in the general case. We experimented also $\alpha = 0.35$ which is closer to repeatability for this kind of image with 3D texture. The value for β was set to $\frac{\alpha}{1020}$ as the data base has 1020 images.

The experiments were conducted in the following way: half of the images from the second view were considered; for each of them a subimage which represents x% of the search image from the first view was taken for the query. x was varied from 100 to 9. The rank of the right answer was measured as the output. As the standard deviation of such answer is high for small window, random selection of such windows were multiplied in order to get a significant mean value.

Four different matching strategies were investigated:

1. simple count of the number of matches;
2. count of the number of matches using the semi-local constraint (see 3.5);
3. Bayesian decision without integrating the unmatched features following therefore the model described by equation 10 and with the two different α's as indicated;
4. Bayesian decision integrating the unmatched features following the model described by equation 11. *Due to technical reasons and lack of time, the results concerning this last case are not integrated in the submitted paper.*

4.2 Results

Fig. 7 displays the behavior of the different strategies. The abcissa represents the size of the subwindows considered for the request (percentage of the image surface). The ordinate shows the mean rank of the correctly matching image.

The four curves displayed correspond, from top to bottom, to the simple voting on invariant, the Bayesian model with $\alpha = 0.5$, the Bayesian model with $\alpha = 0.35$, and the use of semi local context.

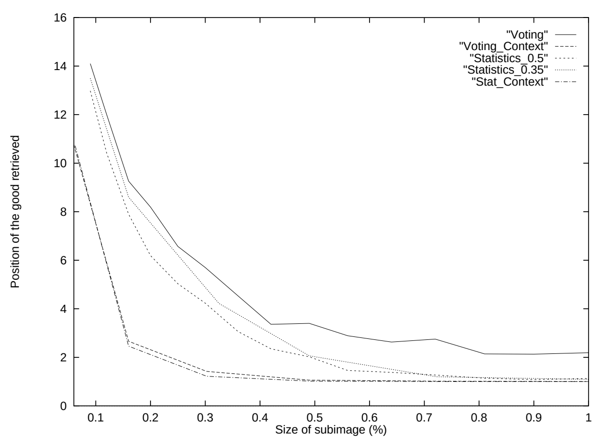

Figure 7: Behavior of voting *vs* posterior probability

The results show that there is an advantage to using the repeatability factor for the kind of image used, but that the behavior of the Bayesian decision rule is not too much affected by this choice. The use of the a priori knowledge distribution offers a clear gain for large windows, by about a factor 2. This gain is more limited for smaller windows where the number of matched features decreases largely. An additional study of the influence of possible outliers on this result is ongoing.

As forcast in the discussion in 3.5, the semi-local constraint improves the results. It is almost perfect for large windows (the right answer is always the first for up to half of the window size) as is the Bayesian decision rule. For very small window, this method outperforms the Bayesian decision rule by a factor 2.

5 Discussion and conclusion

We have presented a Bayesian model for improving the image retrieval procedure using invariant feature vectors. This model has been compared to the previous indexing technique using experimental data set of very similar images, taken in different viewing conditions.

Bayesian decision definitely improves the voting procedure. The gain is substantial when enough features can be matched, and becomes small otherwise. This gain is achieved at the costs of estimating the a priori density of the possible value for the invariant vectors used for indexing the image data base and estimating the repeatability of the point of interest extraction. Both data can be estimated globally in order not to be application dependant. Using a repeatability probability tuned the kind of image used provided only a small improvement. For large query images, the method finds the correct answer almost always first.

One could argue that a factor of 2 in the image retrieval process is not enough for the corresponding cost. One clear reason that the gain is not better is that the invariant used are already very discriminant and the probability that a random match will occur is low. In such a case the counting decision rule becomes very close to the Bayesian decision rule (cf 3.5). This is typically what happens with the use of semi-local constraints: they are very discriminant and almost no false matches occur. More impressive is that for small windows, when few matches are obtained, this method still has the best behavior, almost 50% better than the Bayesian model. It has to be said that comparisons between these two last method is unfair, as they are not working with similar data: the semi local constraint method uses the topological consistency between the extracted points which is not used in the Bayesian decision rule.

More experimental results have still to be conducted in order to get a clearer view of the advantages of this work. Much more important is the construction of a Bayesian decision rule on the top of the semi-local approach in order to improve it.

References

[1] T.O. Binford and T.S. Levitt. Quasi-invariants: Theory and exploitation. In *Proceedings of* DARPA *Image Understanding Workshop*, pages 819–829, 1993.

[2] J.B. Burns, R.S. Weiss, and E.M. Riseman. The non-existence of general-case view-invariants. In J.L. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, chapter 6, pages 120–131. MIT Press, 1992.

[3] L. Florack. *The Syntactical Structure of Scalar Images.* PhD thesis, Universiteit Utrecht, November 1993.

[4] J.J. Koenderink and A.J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

[5] A. Pentland, R.W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, 1996.

[6] Romeny. *Geometry-Driven Diffusion in Computer Vision.* Kluwer Academic Publishers, 1994.

[7] C.A. Rothwell. *Object Recognition Through Invariant Indexing.* Oxford University Press, 1995.

[8] A.H. Salden, B.M. ter Haar Romeny, L.M.J. Florack, M.A. Viergever, and J.J. Koenderink. A complete and irreductible set of local orthogonally invariant feature of 2-dimensional images. In *Proceedings of the 11th International Conference on Pattern Recognition, The Hague, Netherland*, pages 180–184, 1992.

[9] B. Schiele and J.L. Crowley. Object recognition using multidimensional receptive field histograms. In *Proceedings of the 4th European Conference on Computer Vision, Cambridge, England*, pages 610–619, 1996.

[10] B. Schiele and J..L. Crowley. Probabilistic object recognition using multidimensional receptive field histogram. In *Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria*, pages 50–54, 1996.

[11] C. Schmid and R. Mohr. Combining greyvalue invariants with local constraints for object recognition. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, San Francisco, California, USA*, June 1996. ftp://ftp.imag.fr/pub/MOVI/publications/Schmid_cvpr96.ps.gz.

[12] C. Schmid and R. Mohr. Image retrieval using local characterization. In *Proceedings of the IEEE International Conference on Image Processing*, volume II, pages 781–784, September 1996.

[13] H. Schulz-Mirbach. Constructing invariant features by averaging techniques. In *Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, Israel*, pages 387–390, 1994.

[14] M.J. Swain and D.H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

[15] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, USA*, pages 586–591, 1991.

[16] B.M Romeny L.M.J Florack A.H Salden M.A Viergever. Higher order differential structure of images. *Image and Vision Computing*, 12(6):317–325, 1994.