Sparse Representations for Image Decomposition with Occlusions

Mike Donahue [1] Davi Geiger [2] Robert Hummel [2] Tyng-Luh Liu [2]

[1] IMA
University of Minnesota
Minneapolis, MN 55455

[2] Courant Institute
New York University
New York, NY 10012

Abstract

We study the problem of how to detect "interesting objects" appeared in a given image, I. Our approach is to treat it as a function approximation problem based on an over-redundant basis, and also account for occlusions, where the basis superposition principle is no longer valid. Since the basis (a library of image templates) is over-redundant, there are infinitely many ways to decompose I. We are motivated to select a sparse/compact representation of I, and to account for occlusions and noise. We then study a greedy and iterative "weighted L^p Matching Pursuit" strategy, with $0 < p < 1$. We use an L^p result to compute a solution, select the best template, at each stage of the pursuit.

1 Introduction

In the field of signal processing and computer vision an input signal or image is a function f over some subset of $\mathbb{R}$ or $\mathbb{R}^2$. To manipulate and analyze f, it is useful to introduce a linear decomposition into basis elements f_j, i.e., $f = \sum_j c_j f_j$. An example of a well known and useful decomposition of this type is the Fourier series expansion.

We study the object recognition problem via a robust template decomposition approach. Let the image to be recognized be I and the template library be $\mathcal{L}$. The task of image recognition is reduced to a function approximation problem of the form

$$I(x) = \sum_j \sum_i c_{ij} A_i(\tau_j)(x) = \sum_{i,j} c_{ij} T_{ij}(x) \qquad (1)$$

where $\tau_j \in \mathcal{L}$, $T_{ij} = A_i(\tau_j)$ denotes an affine transformation A_i, though in our studies we have just considered translations, applied to the template τ_j, and c_{ij} is the choice of coefficients that "best" decompose the image. Typically the library $\mathcal{L}$ is large, we have an over-redundant basis leading to infinitely many solutions, c_{ij}, to this problem. That is not the case for the Fourier decomposition. The problem is then, to formulate a coefficient selection criterion and a method to compute the coefficients that yields compact representations.

Decomposition with Occlusions: Here we depart from our previous work [14]. To decompose an image containing occluded objects, a special form of (1) is needed. Since occlusions occur at overlapping objects, we introduce an occlusion function O such that

$$\mathrm{O}(\mathrm{T}_{ij})(x) = \begin{cases} T_{ij}(x) & \text{if } T_{ij} \text{ is the toppest one} \\ & \text{among all covering } x \\ 0 & \text{otherwise.} \end{cases}$$

The image decomposition problem with occlusions can be written as

$$I(x) = \sum_{i,j} c_{ij} O(T_{ij})(x) \,. \qquad (2)$$

We will treat O like a polymorphic function that it can also be applied to a pixel x such that

$$O(x) = \begin{cases} 1 & \text{if pixel } x \text{ is covered by some template} \\ 0 & \text{otherwise.} \end{cases}$$

1.1 Coefficient selection and concave optimization

Our approach is to construct an objective function $F(\mathbf{c})$ that when minimized selects a best representation, $\mathbf{c}^*$, among all solutions $\mathbf{c}$ that satisfy the constraint $I(x) = \sum_j \sum_i c_{ij} O(T_{ij})(x)$. We require

1. **Sparse Representation:** represent (decompose) an image using as few templates as possible in order to have an economical (minimal) representation. Field [9] also argued for sparse image representations in the brain.

2. **Occlusions:** allow for partial occlusions, i.e., the cost of fitting a template must take into account that portions of the template may have a "bad match".

3. **Noise:** model noise via "noise templates" accounting for the difference between the template fit and the image. This leads us to search for cost functions that escalate with the magnitude of c_{ij}, but should not dominate the first condition, i.e., the rate of increase in cost as a function of $|c_{ij}|$ should decrease.

The above consideration leads us naturally to adopt concave objective functions. In particular, we will primarily study the objective function

$$F_p(\mathbf{c}) = \sum_{j=1}^{M}\sum_{i=1}^{N} \omega_{ij}|c_{ij}|^p \,, \qquad (3)$$

where N is the number of possible (translations) transformations and M is the size of the template library. The weights ω_{ij} are positive scalars, e.g., they may be set to 1 or to the inverse of the template and image variances.

The sparsity of templates suggests $\mathbf{p} = 0$ to count the number of utilized templates (weighted by ω_{ij}). Noise templates should be paid according to how large the "repair" is, i.e., how large the error c_{ij} is. The balance between both processes, sparsity of the templates and noise modeling leads to values of $0 < \mathbf{p} < 1$. Furthermore, this balance also accounts for occlusions. Compared to conventional L^2 methods ($\mathbf{p} = 2$), $0 < \mathbf{p} < 1$ will cost less for regions where the error is large between the template and the image (occlusions).

The objective function is non-convex, and in fact the optimization problem will generally have multiple local minima. It is possible to characterize all local minima and obtain the global one by visiting them [7]. Since the number of local minima grows exponentially with the size of the template library we consider an alternative greedy algorithm.

Recently, and independently, Chen and Donoho [3, 4] studied the overcomplete signal representation problems with L^1 norm optimization. Their method is based on linear programming, which is efficient, but only applies to the $\mathbf{p} = 1$ case and still leads to a slow algorithm.

Comparison with principal component analysis/Eigenfaces: Our approach is fundamentally different from the "eigenfaces" approach (PCA approach) [17]. In our case the basis functions are fixed and the adaptation of the method is on choosing the appropriate coefficients (from a redundant basis) with the occlusion factor taken into account, a non-linear process. In the PCA approach the choice of basis function, a linear process, is where the adaptation first occurs, and the whole process of choosing coefficients is also linear. PCA works well when the task function is a simple linear superposition of the basis functions. A clear scenario to show how different these methods can be is the case of edge detection. Suppose we have a few different images to train and another image to test the edge detection for both approaches. Our method, would define edges as step edges and then "look" at any of the images (e.g., the test one) to decompose it into these operator and an ok edge detector would have been built. A PCA approach would find the "edge-eigenfunctions" from the training set of images, and hope to describe the next test image by these edge-eigenfunctions (where the edges are not formed from linear superpositions of the training edge-images). It would be a disastrous edge detector! The same should occur for face recognition unless previous "super-normalization" puts them aligned (including emotional expressions normalization).

1.2 Matching Pursuit

Inspired by Mallat and Zhang's work [15] and more recently Bergeaud and Mallat [2], we consider a matching pursuit strategy where, at each stage, the criteria of best selection is based on minimizing an image residue. In regression statistics, this decomposition method is known as *Projection Pursuit Regression* , a non-parametric method that is concerned with "interesting" projections of high dimensional data (see Friedman and Stuetzle [10], Huber [11]).

The original matching pursuit is based on the standard L^2 method. In recognition of image with occlusions, the L^2 norm is not suitable. We propose an L^p matching pursuit, $0 < \mathbf{p} < 1$, to improve the robustness. With $0 < \mathbf{p} < 1$, we lost the structure of inner product but the notion of projection can be recaptured, the criterion for a template to be "best matching" or "closest" to the image is to minimize a cost function. (We will adopt the term "L^p norm" though it is not really a norm.) This modification improves robustness of the pursuit scheme but the convergence of L^p pursuit is now not guaranteed. The energy conservation equation and so Jones' proof [13] of convergence of projection pursuit no longer hold.

2 Template Library and Image Coordinates

We first establish a well-defined over-redundant library of templates containing non-canonical templates as well as one canonical template. A canonical template is a trivial template with zero gray-level value pixels everywhere except one pixel at the extreme left and top corner that its gray-level value is 1. Moreover, we will assume we can apply translations to each template (in theory one could apply any affine transformation with more computational cost). Clearly, this single canonical template plus a set of all translations form a basis for the image space.

Coordinate transformations: Assume that the template library $\mathcal{L} = \{\tau_j : j = 1...M\}$, where we will use $\epsilon_1 \equiv \tau_1$ to represent the canonical template. Let the image to be recognized be I of dimension N and each template τ_j be of dimension N_T (we assume that both N and N_T are perfect square numbers). Furthermore, let $P = \{p_1, p_2, \cdots, p_N\}$ and $Q = \{q_1, q_2, \cdots, q_{N_T}\}$ be the pixel sets of I and any τ_j, respectively. (We order the pixels from top to bottom and left to right.) Let the translation $A_i(\tau_j)$ indicate that the first template pixel q_1 is positioned at the i-th pixel $p_i \in P$ (see Figure 1). We can explicitly describe such relation as follows:

$$Q \xrightarrow[A_i]{} \Omega_i = \{p_k : k \in \Gamma_i \,,\, \Gamma_i \subset \{1, 2, ..., N\}\} \subset P. \quad (4)$$

The mapping formula for A_i is such that $q_r \mapsto p_k = p_{k(r,i)}$ where [1] $k = i + (\lfloor \frac{r-1}{\sqrt{N_T}} \rfloor \times N) + (r - 1 - \lfloor \frac{r-1}{\sqrt{N_T}} \rfloor \times \sqrt{N_T})$. Denote $T_{ij} = A_i(\tau_j)$ and $e_{i1} = T_{i1} = A_i(\epsilon_1)$ [2], then we have $T_{ij}(p_k) = \tau_j(q_r)$. Using these notations, one can write (2) as

$$\begin{aligned} I(p_k) &= \sum_{i=1}^{N} c_{i1} e_{i1}(p_k) + \sum_{j=2}^{M} \sum_{i=1}^{N} c_{ij} O(T_{ij})(p_k) \\ &= \sum_{\lambda=1}^{N} c_\lambda e_\lambda(p_k) + \sum_{\lambda=N+1}^{M.N} c_\lambda O(T_\lambda)(p_k) \qquad (5) \end{aligned}$$

where $\lambda = \lambda(i,j) = (j-1) \times N + i$. We may write $I[k]$, $e_\lambda[k]$ and $O(T_\lambda)[k]$ instead of $I(p_k)$, $e_\lambda(p_k)$ and $O(T_\lambda)(p_k)$, respectively, for simplification.

3 Optimization Problem and Solution

Equation (5) can be written in matrix notation as $\mathrm{Tc} = \mathrm{I}$ where T is

$$\begin{pmatrix} e_1[1] & \cdots & e_N[1] & O(T_{N+1})[1] & \cdots & O(T_{MN})[1] \\ e_1[2] & \cdots & e_N[2] & O(T_{N+1})[2] & \cdots & O(T_{MN})[2] \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ e_1[N] & \cdots & e_N[N] & O(T_{N+1})[N] & \cdots & O(T_{MN})[N] \end{pmatrix},$$

[1]The expression $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.

[2]Note that $e_{i1}(p_j) = e_i(p_j) = \delta_{ij}$, where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ otherwise.

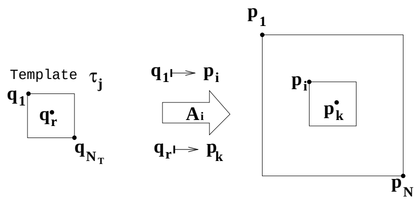


Figure 1: The pixel correspondences between I and $T_\lambda = T_{ij} = A_i(\tau_j)$. We see that pixel q_1 is positioned on p_i and q_r on p_k, respectively.

$$\mathbf{c} = (c_1, c_2, \ldots, c_{MN})^t \quad \text{and} \quad \mathrm{I} = (I[1], I[2], \ldots, I[N])^t \,.$$

Note that if the prototype library forms a basis (linearly independent), then $M = 1$, and there is no freedom in choosing the coefficients $\langle c_\lambda \rangle$; the coefficients are uniquely determined by the constraint. If there are linear dependencies in the prototype library, then $M > 1$, the prototype library over-spans, and the set of all solutions $\langle c_\lambda \rangle$ to the constraint forms an $(M-1)\,N$ dimensional affine subspace in the $M.N$-dimensional coefficient space. Let S denote this solution space, i.e., $\dim(S) = (M-1)\,N$. Using the above matrix notations, our optimization problem can be formulated as:

$$\underset{\mathbf{c}}{Min}\, F_p(\mathbf{c}) = \underset{\mathbf{c}}{Min} \sum_{\lambda=1}^{MN} \omega_\lambda |c_\lambda|^p \qquad \text{subject to } \mathrm{T}\mathbf{c} = \mathrm{I} \qquad (6)$$

where $\mathrm{T} \in \mathbb{R}^{N \times M.N}$, $\mathbf{c} \in \mathbb{R}^{M.N}$, $\mathrm{I} \in \mathbb{R}^N$, $M > 1$. The constraint space, S, is the set of all $\mathbf{c}$ satisfying $\mathrm{T}\mathbf{c} = \mathrm{I}$, and is an affine subspace of dimension $(M-1)\,N$. We will first study the L^p-cost function in (6). It is natural when analyzing F_p in (6) as a function in the coefficient space $\langle c_\lambda \rangle$ to decompose the domain into octants, where each coefficient is of constant sign. This allows the removal of the absolute values in (6), so we may treat F_p as a smooth function inside each octant. For example, if we consider the restriction of F_p to the octant consisting of all points $\mathbf{c}$ such that $c_1 < 0$, $c_2 < 0$, and $c_\lambda > 0$ for $\lambda \geq 3$, then the cost function in (6) becomes

$$F_p(\mathbf{c}) = \omega_1(-c_1)^p + \omega_2(-c_2)^p + \sum_{\lambda=3}^{MN} \omega_\lambda c_\lambda^p \,.$$

Moreover, it is clear that $F_p(\mathbf{c}) \to \infty$ as $||\mathbf{c}|| \to \infty$, so for minimization purposes it suffices to consider bounded $\mathbf{c}$. The bound will depend upon the constraint equation (2), but, for example, if $\mathbf{c}_0$ is any

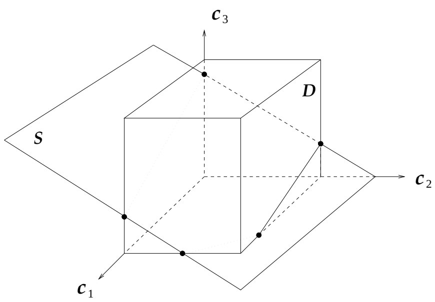


Figure 2: Illustration of a domain restriction polytope obtained from the intersection of a 2 dimensional constraint space S with a rectangular solid bound domain D in a 3 dimensional coefficient space. In this example the intersection is a non-regular pentagon. If the restricted objective function F is concave, then its local minima occur at the vertices of the pentagon.

solution to (2), then it suffices to consider only those $\mathbf{c}$ satisfying $|c_\lambda| \leq (F_p(\mathbf{c}_0)/\omega_\lambda)^{1/p}$ for all λ. Recall that each ω_λ is a positive scalar and can be computed in advance. When combined with the restriction to octants, we have a decomposition of the pertinent domain of F_p into $M.N$-dimensional cubes of edge length $(F_p(\mathbf{c}_0)/\omega_\lambda)^{1/p}$.

The intersection of the constraint space S with these domain cubes gives rise to convex polytopes, as illustrated in Figure 2. The system of domain restrictions can be written out explicitly. For the first (positive) octant they are

$$\begin{aligned} \mathrm{T}\mathbf{c} &= \mathrm{I} \\ c_\lambda &\leq d_\lambda, \qquad 1 \leq \lambda \leq M.N \qquad (7) \\ -c_\lambda &\leq 0, \qquad 1 \leq \lambda \leq M.N, \qquad (8) \end{aligned}$$

where previously we considered the case that each d_λ is at least as large as $(F_p(\mathbf{c}_0)/\omega_\lambda)^{1/p}$.

The relation $c_1 = (1, 0, \ldots, 0)^t \cdot \mathbf{c} \leq d_1$ describes a half-space in the space $\langle \mathbf{c} \rangle$, and the entire collection (7) and (8) together describe the intersection of $2MN$ half-spaces, i.e., a polytope with at most $2MN$ faces. The general inequality defining a half-space is $v \cdot \mathbf{c} \leq d_\lambda$, where v is a vector normal to the bounding hyperplane, and d_λ determines an offset from the origin. So an arbitrary convex polytope having N' faces can be described in the form $\mathrm{B}\mathbf{c} \leq d$, where $\mathrm{B} \in \mathbb{R}^{N' \times M.N}$, $d \in \mathbb{R}^{N'}$, and the inequality is interpreted coordinatewise. So the generalized constraint relations can be written:

$$\begin{aligned} \mathrm{T}\mathbf{c} &= \mathrm{I} \\ \mathrm{B}\mathbf{c} &\leq d. \end{aligned} \qquad (9)$$

The relations (9) can be viewed as defining a polytope inside the affine space S. If we were to perform a basis transformation to obtain coordinates conducive to representations inside S, then F_p under the same transformation would loose its simple form. Even without this consideration, it is useful to study more general objective functions. The specific property of F_p of interest to us is *concavity.* A function F mapping from a convex domain Ω of a vector space X to $\mathbb{R}$ is concave if

$$F(\alpha x + (1-\alpha)y) \geq \alpha F(x) + (1-\alpha)F(y)$$

for all x and y in Ω and $\alpha \in [0,1]$. The result we desire (Proposition 1) actually requires only a weaker property, which we call *pseudo-concave.* A function $F : \Omega \to \mathbb{R}$ as above is pseudo-concave if

$$F(\alpha x + (1-\alpha)y) \geq Min\{F(x), F(y)\}$$

for all x and y in Ω and $\alpha \in [0,1]$. Clearly any concave function is also pseudo-concave.

Proposition 1 *Let Ω be a closed, bounded, convex polytope in a vector space X, and let $F : \Omega \to \mathbb{R}$ be pseudo-concave. Then the global minimum of F on Ω occurs at a vertex of Ω.*

4 Multiple Templates and Matching Pursuit

Let us assume that the residue at the initial stage is the input image, i.e., $R^0 I = I$. Then, at stage n, if a transformed template $T_{\lambda_n} (= T_{i_n j_n} = A_{i_n}(\tau_{j_n}))$ and coefficient c_{λ_n} are chosen, the n-th residual image can be updated by "projecting" the $R^{n-1}I$ in the direction of T_{λ_n}. More precisely,

$$R^n I[k] = R^{n-1} I[k] - c_{\lambda_n} T_{\lambda_n}[k](1 - O[k]), \qquad (10)$$

where $k = 1...N$ and $\lambda_n = (j_n - 1) \times N + i_n$. Note that T_{λ_n} is only of dimension N_T and we have assumed that $T_{\lambda_n}[k] = 0$ if $k \notin \Gamma_{i_n}$ as defined in (4).

Let Γ be the index set of matched pixels (those covered by the selected transformed templates in the process of L^p matching pursuit). Clearly, $\Gamma = \emptyset$ initially and is updated, say at stage n, as $\Gamma \leftarrow \Gamma \cup \Gamma_{i_n}$. Suppose that L^p matching pursuit is completed in m stages and T_{λ_n} is the best selected template with an associated index set Γ_{i_n} at stage n, $1 \leq n \leq m$. Then, we have $\Gamma = \bigcup_{n=1}^{m} \Gamma_{i_n}$ and

$$I(x) = \sum_{n=1}^{m} c_{\lambda_n} O(T_{\lambda_n})(x) + \sum_{k=1}^{N} c_k e_k(x) \qquad (11)$$

$$= \sum_{n=1}^{m} c_{\lambda_n} O(T_{\lambda_n})(x) + \sum_{k\in\Gamma} c_k e_k(x) + \sum_{k\notin\Gamma} I[k] e_k(x)\,,$$

where $\sum_{n=1}^{m} c_{\lambda_n} O(T_{\lambda_n})$ is the main decomposition, and $\sum_{k\in\Gamma} c_k e_k$ is the residual decomposition. The decomposition cost F_p for (11) is

$$\sum_{n=1}^{m} \omega_{\lambda_n} |c_{\lambda_n}|^p + \sum_{k\in\Gamma} \omega_k |c_k|^p + \sum_{k\notin\Gamma} \omega_k |I[k]|^p =$$

$$\sum_{n=1}^{m} \omega_{\lambda_n} |c_{\lambda_n}|^p + \sum_{k\in\Gamma} (\omega_k |r_k|^p - |I[k]|^p) + \sum_{k\in\{1..N\}} |I[k]|^p \qquad (12)$$

The weights related to selection of canonical template are defined as $\omega_k = 1$ if $k \notin \Gamma$ or $\omega_k = \omega_{\lambda_n}$ if $k \in \Gamma$ and T_{λ_n} is the toppest one covering p_k. For every possible decomposition of I, the set Γ in (11) could be different but the last term in (12) is common for all decompositions.

The cost F_p in (6) and (12) are total cost. We need to formulate a stage-wise cost function, denoted as $\mathcal{F}_p$, for the pursuit. Note that $\mathcal{F}_p$ is a function of one c (a scalar) and we have also approximated the computation of $\mathcal{F}_p$ to exclude the overlap region and only after a best matching is decided we resolve overlapping ambiguities if there are any. Intuitively, one may define $\mathcal{F}_p^n$ for selecting T_{λ_n} with coefficient c_{λ_n}, at stage n, as

$$\omega_{\lambda_n} |c_{\lambda_n}|^p + \sum_{k\in\Gamma_{i_n}} (\omega_{\lambda_n} |r_k|^p - |I[k]|^p)(1 - O[k])\,, \qquad (13)$$

where $r_k = |R^{n-1} I[k] - c_{\lambda_n} T_{\lambda_n}[k]|$ is the residue at p_k. Our experiments show that the cost (13) may prefer to match white regions (0 is black, 255 is white). In stead, we define $\mathcal{F}_p^n(c_{\lambda_n})$ as

$$\frac{1}{\beta_{\lambda_n}} \left(\omega_{\lambda_n} + \sum_{k\in\Gamma_{i_n}} \frac{(\omega_{\lambda_n} |r_k|^p - |I[k]|^p)}{|c_{\lambda_n}|^p} (1 - O[k]) \right)\,, \qquad (14)$$

where β_{λ_n} is the number of pixels covered by T_{λ_n} with $O[k] = 0$ and $\omega_{\lambda_n} = 1/|Variance(\tau_{j_n})|^{\frac{p}{2}}$. Our experiment results show by adapting (14) as the stage-wise cost function, the pursuit can avoid "over-utilization" of templates on dark or white regions, since it is possible for a template to match very well in a darker region due to a small value of $|c_{\lambda_n}|$. The weight ω_{λ_n} give the benefits of choosing a template with more variation (large variance) over a plain one.

5 Matching Pursuit Experiments

Synthetically Randomized Images : Let's begin with a simple but instructive experiment to test our template matching algorithm for a synthetic example. In this experiment, the template library $\mathcal{L}$ consists of three different types (or shapes) of templates ((a), (b), (c) in Figure 3). There are 40 templates for each type so that $\mathcal{L}$ includes 120 non-canonical templates and one canonical template ϵ_1. Each of the non-canonical template is a synthetically randomized image with gray-level values between $(0, 200)$ generating from a random number generator. To construct a test image I_1 (as in Figure 3-(d)), we first select one non-canonical template randomly from each template type in $\mathcal{L}$ to form the base (exact) image then add noise and an occluded square derived from uniform distribution in $(0, 10)$ and $(245, 255)$, respectively. The threshold values used in simulation vary with respect to the value of $\mathbf{p}$ for L^p matching pursuit. Our results suggest for $\mathbf{p} \in (0.25, 0.75)$, the L^p pursuit is rather robust. But, as shown in Figure 3-(f) R_2, it failed to recognize the occluded object for $\mathbf{p} \geq 0.75$.

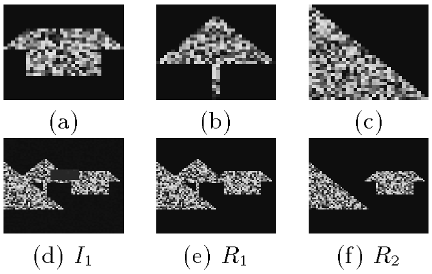


Figure 3: (a), (b), (c) are synthetic template type 1, type 2 and type 3, respectively. (d) Test image I_1 with noise added and occlusion. (e) Result of the decomposition for the L^p with $\mathbf{p} = 0.25$. (f) Results once the breakdown limits are reached, and occluded objects are not recognized when $\mathbf{p} \geq 0.75$.

Face Recognition : A small library of face templates has been established (see Figure 4 (a)-(j)). The dimension of all 10 face templates is 64×64. Numerous experiments have been carried out to test our algorithm. To illustrate, consider the two real images, I_1 and I_2, in Figure 5 (a)-(b). We obtained decomposition results R_1 and R_2 shown in Figure 5, for $\mathbf{p} = 0.25$. (Similar results are derived for $\mathbf{p} = 0.50,\ 0.75$.) When $\mathbf{p} = 2$, it is indeed the L^2 matching pursuit method and the recognition results are R_3 and R_4. Our proposed L^p matching pursuit has the robustness advantage over the L^2 one.

References

[1] J. Ben-Arie and K. R. Rao, *"On the Recognition of Occluded Shapes and Generic Faces Using Multiple-Template Expansion Matching"*, Proceedings IEEE International Conference on Pattern Recognition, New York City, 1993.

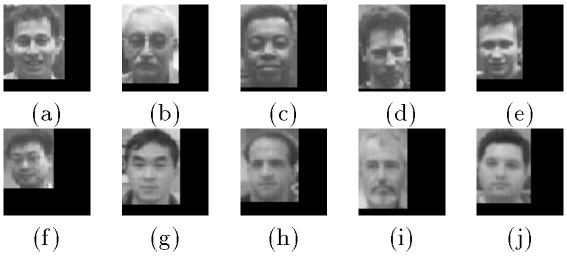
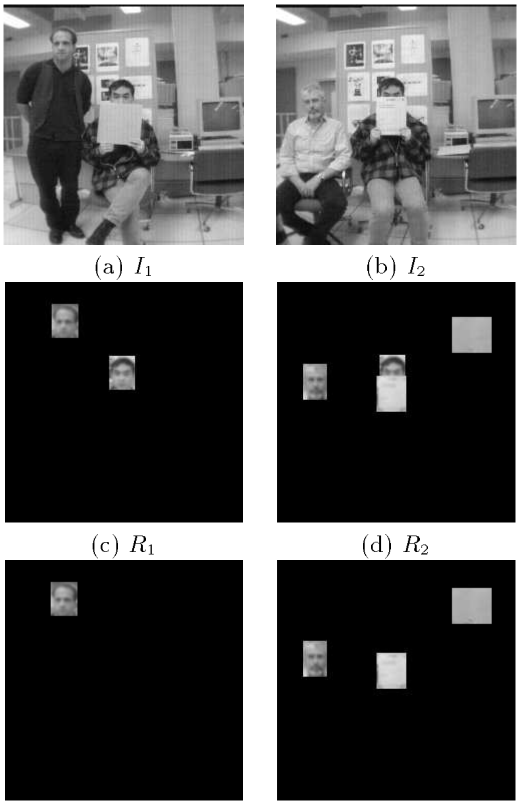


Figure 4: (a) - (j) are the 10 face templates used in the face recognition experiment.



(a) I_1 (b) I_2

(c) R_1 (d) R_2

(e) R_3 (f) R_4

Figure 5: (a),(b) The test images, where some templates are present with **small** distortions (scale and viewing angle), noise and occlusions. (c),(d) Image decomposition for L^p matching pursuit with $\mathbf{p} = 0.25$ (similar results are obtained for p up to 0.75). (e),(f) Image decomposition for $\mathbf{p} = 2.0$ and recognition is destroyed (this is equivalent to use correlations methods, like in the L^2 matching pursuit). Note that in decomposition for I_2, we have used two booklike templates besides the face templates.

[2] F. Bergeaud and S. Mallat, *"Matching Pursuit of Images"*, SPIE, Orlando, 1995.

[3] S. Chen and D. Donoho, *"Atomic Decomposition by Basis Pursuit"*, Technical Report, Stanford University, May, 1995.

[4] S. Chen and D. Donoho, *"Basis Pursuit"*, TR, Stanford University, Nov. 1994.

[5] R. Coifman and V. Wickerhauser, *"Entropy-based Algorithms for Best Basis Selection"*, IEEE Transactions on Information Theory, vol. 38, no. 2, 1992.

[6] T. H. Cormen, C. E. Leiserson and R. L. Rivest, *Introduction to Algorithms*, McGraw-Hill, 1990.

[7] M. J. Donahue and D. Geiger, *"Template Matching and Function Decomposition Using Non-Minimal Spanning Sets"*, Technical Report, Siemens, 1993.

[8] H. Ekblom, *"L_p-methods for Robust Regression"*, BIT 14, p.22-32, 1973.

[9] D. Field, *"What Is the Goal of Sensory Coding"*, Neural Comp. 6, p.559-601, 1994.

[10] J. H. Friedman and W. Stuetzle, *"Projection Pursuit Regression"*, Journal of the American Statistical Association, vol. 76, p.817-823, 1981.

[11] P. J. Huber, *"Projection Pursuit"*, The Ann. of Stat., vol. 13, No.2, p.435-475, 1985.

[12] P. J. Huber, *Robust Statistics*, John Wiley & Sons, New York, 1981.

[13] L. K. Jones, *"On a Conjecture of Huber Concerning the Convergence of Projection Pursuit Regression"*, The Ann. of Stat., vol. 15, No.2, p.880-882, 1987.

[14] T. Liu, M. Donahue, D. Geiger and R. Hummel, *Image Recognition with Occlusions*, Proc. 4th European Conf. on Computer Vision, London, 1996.

[15] S. Mallat and Z. Zhang, *"Matching Pursuit with Time-Frequency Dictionaries"*, IEEE Trans. on Signal Processing, Dec. 1993.

[16] P. J. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*, John Wiley, New York, 1987.

[17] M. Turk and A. Pentland, *"Eigenfaces for Recognition"*, J. of Cognitive Neuroscience, vol. 3, p.71-86, 1991.

[18] B. Uhrin, *"An Elementary Constructive Approach to Discrete Linear l_p-approximation, $0 < p \leq +\infty$"*, Colloquia Mathematica Societatis János Bolyai, 58. Approximation Theory, Kecskemét, 1990.