PROBABILISTIC LINEAR MODELS FOR MULTIRESOLUTION ESTIMATION IN GRAY-LEVEL IMAGES

José MARTÍNEZ-AROZA
Dept. de Matemática Aplicada,
Universidad de Granada, SPAIN.

Ramón ROMÁN-ROLDÁN, member IEEE
Dept. de Física Aplicada,
Universidad de Granada, SPAIN.

KEY PHRASES: gray-level image, histogram estimation, probabilistic models for images, multiresolution, image processing, nonlinear filter

All correspondence, including galley proofs, should be addressed to

Ramón ROMÁN-ROLDÁN

Dept. de Física Aplicada.
Facultad de Ciencias, Universidad de Granada.

Avda. Fuente Nueva, s/n.
18071 Granada, SPAIN

Tel (34) 58 24 32 22, Fax (34) 58 24 32 14,
E-mail **rroman@ugr.es**

PROBABILISTIC LINEAR MODELS FOR MULTIRESOLUTION ESTIMATION IN GRAY-LEVEL IMAGES

José MARTÍNEZ-AROZA
Ramón ROMÁN-ROLDÁN, member IEEE*

KEY PHRASES: gray-level image, histogram estimation, probabilistic models for images, multiresolution, image processing, nonlinear filter

Abstract

A multiresolution analysis of digital gray-level images is presented. A gray-level multi-scale framework is determined from two main assumptions: the gray scale is binary at the finest spatial resolution, and the gray levels of composed regions are obtained additively. In order to interrelate the gray-level histograms of the same image at different resolutions, probabilistic linear models are developed, which are then applied for estimation. Linear-optimization theory is used as a way of constructing such models. A general procedure for image processing is sketched, based on gray-level estimation. A versatile algorithm for nonlinear filtering is derived. Some examples of prospective applications are given.

*This work was partially supported by grant TIC91-646 from the DGYCIT of the Spanish Goverment.

I INTRODUCTION

I.1 BACKGROUND

During the last decade there has been a considerable surge of interest in multiresolution analysis for image processing. Pyramid methods, for example, have commanded considerable attention [1, 2, 3, 5, 6, 7, 8, 20]. The multiresolution concept itself has been worked out from several points of view [9, 10, 22].

The present authors have been working for a few years in a particular line of research within the field of multiresolution analysis of digital images [12, 18, 19]. Applications in image processing, namely nonlinear filtering, are beginning to be developed [15]. For the sake of convenience the main body of these ideas is included within the paper.

Digital gray-level images are generally described and processed with the same gray scale, typically 256-ranged, for all resolutions. This is because both theoretic and computational treatments are easier and the description is more uniform, while in practice there is no need to change the scale. Sometimes, however, the original scene has already a discrete, even binary, set of gray levels. For example, coherently-imaged rough surfaces with speckles, as well as printed images, are cases in which the image is binary at a spatial resolution fine enough. But, because of either the equipment being used, or the adjusting parameters, or any other conditions of observation, an image at a coarser spatial resolution is obtained. Many different gray levels can then be distinguished, depending on the number of black and white pixels included in the minimum region observed. Thus, the original binary gray scale must be expanded up to a level determined by the spatial resolution used to observe the scene.

The basic concern of the present paper is to derive relationships for digital gray-level images described at different spatial resolutions with different corresponding discrete gray scales. In other words, a framework of spatial resolutions near from that of the binary gray scale is selected, whereas the common framework is far from it. The present authors have already explored this subject. Certain relationships between the entropies of the gray-level histograms of a given image at different resolutions have already been studied [19]. The interest of this study lies in the fact that the Shannon entropy of the histogram is a measure of the information conveyed by the image (when considered as a message) under certain assumptions. Also, the problems concerning the estimation of the histogram at a given resolution from the histogram at a coarser resolution have been addressed previously by the present authors [13, 14].

I.2 CONTENT OF THE PAPER

In Section II we first introduce the interdependence between the gray-level histograms at different resolutions. Then, the probabilistic histogram estimation is formalized by means of linear models, which are specific to classes of images. Both elemental and composed linear models are established. Section III shows how linear models can be obtained by solving linear optimization problems, using appropriate constraints. Solutions for some general problems are illustrated. Two interesting cases in particular are posed: one for a non-elemental model leading to the hypergeometric model; another for a nonlinear optimization problem — that of maximizing the resulting entropy. Section IV deals with the subject of nonlinear processing from those models of estimation and applies the theory to design algorithms for image processing. Finally, examples of filtering and superresolution operations are offered.

I.3 DEFINITIONS AND NOTATIONS

When a surface is observed, an image is obtained. The spatial resolution of this image is given by the size of the smallest observable region on the surface. If a second image of the same surface can be obtained with a finer resolution, then the old minimum details can be decomposed into new smaller observable regions. Obviously this decomposition may not always be done within an exact number of regions, but this can be made possible by selecting the appropriate resolutions. Thus, we select these resolutions in order to have a sequence of sizes in powers of two.

The term pixel is used here as the unit for measuring the size of the regions. Although the actual size is merely a matter of suitability in fixing a unit, it is better to think of the pixel as the minimum region at the finest resolution.

The notation concerning spatial resolution is:

$m = 0, 1, 2, \ldots$ the resolution index (from fine to coarse)

m-region — the minimum observable region at resolution m.

$R_m = 2^m$ — the size in pixels of every m-region.

$r_{m;m'} = \frac{R_{m'}}{R_m}$ — quotient either of sizes or total number of regions for the pair $(m; m')$.

Because pixels are defined as black or white, the gray-level scale is binary at this finest resolution. Since regions are composed of pixels, the gray level of a region is its number of black pixels. This actually is equivalent to the natural assumption of gray-level additivity when composing regions, as it happens in the physical domain for the intensity in a noncoherent integration. Therefore both the gray scale and the gray-level resolution depend on the spatial resolution, so that a non-normalized gray scale is used.

The notation concerning the gray-level resolution is:

$K_m = \{0; 1; \ldots; R_m\}$ the gray scale at resolution m.

$k \in K_m$ — the gray level of an m-region, given by the number of black pixels included.

$(m; k)$-region — an m-region with a gray level equal to k.

Figure 1 illustrates all these concepts.

Because of the gray-level additivity, for all m'-regions composed of a partition consisting of $r_{m;m'}$ disjoint m-regions[1], with gray levels $k_i;\ i = 1; \ldots; r_{m;m'}$

$$k' = \sum_{i=1}^{r_{m;m'}} k_i; \qquad m' > m:$$

By counting the number of m-regions for each gray level throughout the image, relative frequencies of gray levels result. They constitute the m-histogram of the image. We shall represent it by a probability distribution. Related notations are:

$N_{m;k}; N_m$ — number of $(m; k)$-regions, m-regions, in the whole image.

$P_{m;k} = \frac{N_{m;k}}{N_m}$ — relative frequency of gray level k, taken as a probability.

$P_m = \{P_{m;k}\}_{k \in K_m}$ the gray-level m-histogram.

[1] The notation about resolutions and gray levels has been slightly changed in this paper with respect to others from the same authors, to improve the readability.

II HISTOGRAM ESTIMATION

The two subjects presented here — the interdependence between histograms of the same image at different resolutions and the estimation of one histogram from the other—, are conceptually distinct, but formalized by the same relationship.

II.1 HISTOGRAM INTERDEPENDENCE

Let us consider two fixed resolutions m, m' with $m < m'$. Let $P_{m'}$ be the m'-histogram of any image given at resolution m'. Even if we cannot obtain the same image at the finer resolution m, it may be still possible to infer something about its m-histogram. The average gray level of the image must obviously be an invariant value in order to achieve compatibility between the corresponding histograms when we change the resolution. This is not sufficient, however, for determining uniquely the m-histogram. Thus, the interdependence between histograms at different resolutions is of key interest. In particular, the relationships [11]

$$P_{m;k} = \sum_{k' \in \mathbf{K}_{m'}} Q(m;k|m',k') P_{m';k'}; \qquad (k \in K_m) \tag{1}$$

answer in a certain way the interdependence question. Here, $Q(m;k|m',k')$ stands for the proportion of the $(m;k)$-regions with respect to the number of m-regions, all within the (m',k')-regions. Each term in the sum of (1) clearly represents the proportion of the $(m;k)$-regions coming from the (m',k')-regions. In matrix notation:

$$P_m = Q_{m;m'} P_{m'} \tag{2}$$

where $Q_{m;m'}$ stands for the $(R_m + 1) \times (R_{m'} + 1)$ matrix

$$Q_{m;m'} = \begin{pmatrix} Q(m;0|m',0) & Q(m;0|m',1) & \cdots & Q(m;0|m',R_{m'}) \\ Q(m;1|m',0) & Q(m;1|m',1) & \cdots & Q(m;1|m',R_{m'}) \\ \vdots & \vdots & & \vdots \\ Q(m;R_m|m',0) & Q(m;R_m|m',1) & \cdots & Q(m;R_m|m',R_{m'}) \end{pmatrix} \tag{3}$$

The proportions Q can be used to define a class of images as being the set of images with the same matrix $Q_{m;m'}$. The set of all images can be divided into classes in this way. This classification procedure of images permits us to state that when we have a matrix $Q_{m;m'}$ arising from a given real image, then we have a representation for a specific class of images, including the one given, and all those that differ in only certain spatial permutations of pixels and regions.

II.2 THE LINEAR MODEL OF ESTIMATION

If only $P_{m'}$ and $Q_{m;m'}$ of an image are known, then its unknown m-histogram can be obtained from the relationships (2), where the Q's take on a wider probabilistic meaning than merely being thought of as frequencies. In fact they can be obtained [18] without counting regions in the image — that is, from certain prior information about either the image or, better still, its class. According to (2), $Q_{m;m'}$ expresses our knowledge about a class of images from the standpoint of the histogram interdependence. In this sense, the matrix $Q_{m;m'}$ will be called a linear model of dependence — due to the linearity of (1)— , of which the elements $Q(m;k|m',k')$ are conditional probabilities.

The class to which a given image belongs may be unknown (the image can be obtainable only at resolution m', not m). Nevertheless, if it can be assigned to a certain model $Q_{m;m'}$, then the m-histogram can be estimated by (2), where $Q_{m;m'}$ is now considered as a linear model of estimation. A good estimation relies exclusively on a good model, i.e. a good approximation of the actual set of conditional probabilities for the proper class of images.

Criteria for setting up models

A key matter is how to find the proper model for a specific kind of problem, not for a particular image. Two different approaches may be followed, the theoretical and the empirical one. In the first, the model is determined from the problem at hand by the general and appropriate features, which usually refer to the class of images for the model to be applied. In the empirical approach, a set of images are considered as a class based on their look or other information; then, they are taken as a training set and processed to get the corresponding empirical model. Some work has been published by the present authors [19] about the theoretical approach; the empirical one, while being still a matter of an ongoing research, will not be treated here.

The theoretical criterion to be satisfied by a desired model may be specified at three different levels:

The column level: A criterion is stablished separately for each column of each matrix ($k', m; m'$ given).

The matrix level: A criterion is tablished for each matrix ($m; m'$ given), the same for all its columns.

The general level: A unique criterion should be satisfied for all columns at all pairs of resolution. One example will be shown in Section II.2.1 as the hypergeometric model, which satisfies the equiprobabilistic composition assumption. At this level, the transitivity between matrices $Q_{m;m'}$ is a desirable property.

Except for the following paragraph, the rest of this section and the main body of the next refer to the column level.

II.2.1 Non-consistent models.

It is possible to consider non-consistent models as a extension of the theory of linear models, the main use for which might be to design processing algorithms, instead of merely estimating histograms. Here we lay out a generalized theory.

A general linear model takes the form (2), in which $P_{m'}$ is the histogram of an input image, I_{IN}, while P_m is the histogram of an output image, I_{OUT}. The case $m = m'$ is allowed as well as the case $m > m'$. A general linear model is intended to estimate P_m from $P_{m'}$.

In this context, if the constraint $I_{IN} = I_{OUT}$ is imposed, then we have consistent models just as they are studied in this paper. Depending on the imposed relationship between I_{IN} and I_{OUT}, different theories will emerge concerning linear models.

Elemental models

For a given pair of resolutions $(m; m'); m < m'$, and for a fixed $k' \in K_{m'}$, the corresponding k'^{th} column of the $(m; m')$-matrix in (3) expresses the average proportions of the $(m; k)$-regions within the $(m'; k')$-regions, which can be composed of m-regions in all possible ways. Let an image be thought as having all the (m', k')-regions divided into (m, k)-regions in the same way. The simple model associated with this special image will be called an elemental model. We shall study these, keeping in mind that the general situation is actually a mixture of elemental models. More specifically, we shall prove that the probability distribution $\{Q(m, k|m', k')\}_{k \in \mathbf{K}_m}$ in the general case is a linear convex combination of all the probability distributions, each associated to a fixed composition.

The number of different compositions, given by combinatorial relationships, has been referred to as compositions with constraints [21]; this is the number of ways a positive integer k' can be obtained as the sum of $r_{m,m'}$ integers of the form $k \in K_m$:

$$\Phi(m, m', k') = \sum_{j=0}^{\left\lfloor \frac{k'}{R_m+1} \right\rfloor} (-1)^j \binom{r_{m,m'}}{j} \binom{k' + r_{m,m'} - 1 - j(R_m+1)}{r_{m,m'} - 1}.$$

Let us consider a particular composition

$$k' = k_1 + k_2 + \cdots + k_{r_{m,m'}}, \quad (k_j \in K_m \ \forall j), \tag{4}$$

of which the associated probability distribution is

$$Q(m, k|m', k') = \frac{n_k}{r_{m,m'}}, \qquad (k \in K_m),$$

where n_k denotes the number of occurrences of integer k in (4). All the probabilities are multiples of $1/r_{m,m'}$, and the model thus constructed is an elemental model.

Composed models

Returning to the general case, assume that all the (m', k')-regions in the given image have been examined, and that the ways in which these regions are composed by m-regions have been ordered from 1 to $\Phi(m, m', k')$. Let λ_i be the proportion of (m', k')-regions divided according to the i^{th} composition, and let $\{Q_i(m, k|m', k')\}$ be the associated probability distribution. Then, for the subset of (m', k')-regions of the image and the class of images represented by this subset, we find the following composed model:

$$Q(m;k|m',k') = \sum_{i=1}^{\Phi(m;m',k')} \lambda_i Q_i(m;k|m',k'); \qquad (k \in K_m)$$

that is

$$Q = \sum_{i=1}^{\Phi(m;m',k')} \lambda_i Q_i; \qquad (0 \le \lambda_i \le 1); \qquad \sum_{i=1}^{\Phi(m;m',k')} \lambda_i = 1; \qquad (5)$$

which is a linear convex combination of the elemental models Q_i.

Because $\Phi(m;m',k')$ is a finite number, we have the following property: "for any linear model, the probability distribution $\{Q(m;k|m',k')\}$ is a point belonging to the convex polyhedron generated in the space R^{R_m+1} by the elemental probability distributions $\{Q_i(m;k|m',k')\}_{i=1}^{\Phi(m;m',k')}$, each one associated to a fixed composition (4)".

The neutral model

As we have seen above (Section II.2), composed models can be built up as linear convex combinations from elemental models. Although these models do not properly optimize any linear objective function, they can be of interest for certain situations. An example is shown below.

Taking into account that many different arrangements of binary pixels may lead to the same composition of an (m',k')-region with m-regions, it is natural to choose the set of weights $\{\lambda_i\}$ in (5) as the set of normalized numbers $\{C_i / \sum C_i\}$ of arrangements possible for each composition i. This combinatorial problem is easy to solve. Let Q_i be the i^{th} elemental model $(1 \le i \le \Phi(m;m',k'))$, which corresponds to the i^{th} composition $k' = k_1 + k_2 + \cdots + k_r$. Let C_i be the number of pixel arrangements making up this composition. This number is

$$C_i = \frac{r_{m;m'}!}{\prod_{k \in K_m} n_{ik}!} \prod_{k \in K_m} \binom{R_m}{k}^{n_{ik}}$$

where n_{ik} stands for the number of occurrences of term k in the i^{th} composition.

Thus, by composing a model via a linear convex combination with coefficients $\lambda_i = C_i / \sum C_i$, we obtain the model [11]

$$Q(m;k|m';k') = \frac{\binom{k'}{k}\binom{R_{m'}-k'}{R_m-k}}{\binom{R_{m'}}{R_m}}; \quad \begin{matrix} m < m' \\ \max\{0; R_m - R_{m'} + k'\} \le k \le \min\{k'; R_m\} \end{matrix}; \qquad (6)$$

which is the well-known hypergeometric distribution.

The following comments are relevant to the hypergeometric solution:

1. The hypergeometric model (Fig. 3e) is similar to the low-contrast one in Fig. 3a, although the first one is not as closely grouped. Thus, it is another conservative model. Its matrix has the same sum $(R_{m'} + 1)/(R_m + 1)$ in all rows, and therefore leads to a uniform m-histogram starting from a uniform m'-histogram.

2. Such a solution is obtained by assigning a weight to each elemental distribution as given by the number of arrangements providing the corresponding composition. Hence, we are assuming that all the arrangements of the pixels in the $(m';k')$-regions are equally probable. This is the most unbiased hypothesis when nothing is known about the m-histogram (except that it must also coexist with the input m'-histogram), in accordance with the Maximum-Entropy Principle.

III LINEAR MODELS AS SOLUTION OF OPTIMIZATION PROBLEMS

This section developes a method to obtain linear models verifying certain ideal conditions. These theoretic models can be applied when dealing with images satisfying the same conditions (or near from them). In the sequel we consider a given environment $(m;m')$, that is, a fixed pair of resolutions $(m;m')$ with $m < m'$.

Para muchas aplicaciones practicas, no hay que tragarse esto, pero para ciertos analisis teoricos, esta seccion tiene un interes intrinseco, y su importancia reside en que simplifica notablemente la busqueda de modelos teoricos.

As demonstrated above, every column of a linear model (3) is a point belonging to a certain convex polyhedron. This suggests the possibility of searching for such a column by optimizing a given linear function of the probability distribution $\{Q(m;k|m',k')\}_{k\in K_m}$, for the proper $k' \in K_{m'}$. In other words, we are solving linear optimization problems (LOP's) in order to find linear models to match some kind of criterion of optimality. This problem can be posed as follows:

Linear Optimization Problem (LOP): Let m', m and k' be fixed. Given a linear function of the m-probability distribution

$$\varphi = \varphi(p_0;p_1;\ldots;p_{R_m}) = \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{R_m} p_{R_m}; \qquad (7)$$

where the α_k are real numbers, find $\{p_0;p_1;\ldots;p_{R_m}\}$ maximizing φ, subject to the (linear) constraints given by the convex polyhedron associated to the environment $(m;m',k')$.

To belong to the convex polyhedron implies a set of constraints, including being a probability distribution, avoiding certain impossibilities ($p_k = 0$ if $k > k'$ or if $R_m - k > R_{m'} - k'$), and preserving the average gray level in (m',k')-regions ($\sum_{k\in K_m} k p_k = k' \frac{R_m}{R_{m'}}$); these constraints, however, are not enough to determine the polyhedron (a study of consistency can be seen in [11]).

Since the solution for this LOP will in any case be a proper vertex of the polyhedron, that is, an elemental model, it is possible to put forward another property: all the p_k in the solution are rational numbers such as

$$p_k = \frac{n_k}{r_{m;m'}}$$

where the n_k are integers. This comes from considering that elemental models arise from fixed compositions of (m',k')-regions with m-regions.

Because some objective functions φ can have helpful meanings, it is useful to study unisolvency (existence and uniqueness of a solution) of the LOP for some kinds of objective functions.

III.1 UNISOLVENCY OF THE LOP FOR SOME OBJECTIVE FUNCTIONS

Given an objective function φ, the associated LOP does not generally have a single solution, since this may be a full edge, face or hyperface of the polyhedron. It may seem necessary to study each separate objective function to reach a conclusion concerning the unisolvency of the LOP. It is possible, however, to state some theoretic results for unisolvency, which hold for certain sets of objective functions with a common characteristic.

Let m' and m be two fixed resolution levels ($m < m'$) and let $k' \in K_{m'}$ be a given gray level ($0 \le k' \le R_{m'}$). Let φ be an objective function given by its coefficients according to (7).

Sequences. φ has a $\cup$-convex, plane or $\cap$-convex sequence in its k^{th} term ($1 \le k \le R_m - 1$), when the quantity $\alpha_{k-1} - 2\alpha_k + \alpha_{k+1}$ is positive, zero, or negative, respectively.

Let us say that an objective function is totally $\cup$-convex, plane or $\cap$-convex when it has $\cup$-convex, plane or $\cap$-convex sequences, respectively, in all its terms from $k = 1$ to $k = R_m - 1$ (see Fig. 2 a,b,c).

Interchanges. Let us consider an (m', k')-region composed of $r_{m,m'}$ m-regions, and having a given distribution of (elemental) black and white pixels. An interchange consists of selecting two pixels, a black one and a white one, in the (m', k')-region, and interchanging their positions. We study interchanges involving different m-regions inside the same m'-region, because such interchanges modify the gray levels of these m-regions. Below, we shall speak only about this kind of interchange.

An interchange will be known as a smoothing interchange (SMI) when the absolute difference of gray levels of the involved m-regions is smaller after the interchange, and it will be called a sharpening interchange (SHI) when this difference is greater.

Any interchange within an (m', k')-region will leave its gray level (and thus the m'-histogram of the image) unchanged, but the m-histogram may differ. Now we shall analyze the effects of an interchange on the value of an objective function.

III.2 RESULTS FOR TOTALLY CONVEX OBJECTIVE FUNCTIONS

Theorem 1 Let us consider an interchange within an (m', k')-region from a given image. Let φ be an objective function, the values of which on the m-histogram of the m'-region are φ_B and φ_A, respectively, before and after interchanging. Thus we have the following results (see Fig. 2, a, b and c):

If φ is	and the interchange is	then
Totally plane	any	$\varphi_A = \varphi_B$
Totally $\cap$-convex	SMI	$\varphi_A > \varphi_B$
Totally $\cap$-convex	SHI	$\varphi_A < \varphi_B$
Totally $\cup$-convex	SMI	$\varphi_A < \varphi_B$
Totally $\cup$-convex	SHI	$\varphi_A > \varphi_B$

Proof. Only the second case is proved (φ totally $\cap$-convex and SMI), the remaining cases being completely analogous. Let $\{p_0, p_1, \ldots, p_{R_m}\}$ be the m-histogram of the m'-region — considered itself as an image— before interchanging. We have $\varphi_B = \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{R_m} p_{R_m}$. Let k_1 and k_2 ($k_1 < k_2$) be the gray levels of the m-regions involved in the interchange. A SMI is possible only if $k_1 + 1 \leq k_2 - 1$. After interchanging, k_1 has increased by 1, and k_2 has decreased by 1. Thus, the new m-histogram $\{q_0, q_1, \ldots, q_{R_m}\}$ shows some alterations with respect to the previous one:

$$\begin{aligned} q_{k_1} &= p_{k_1} - \omega \\ q_{k_1+1} &= p_{k_1+1} + \omega \\ q_{k_2-1} &= p_{k_2-1} + \omega \\ q_{k_2} &= p_{k_2} - \omega \end{aligned}$$

when $k_1 + 1 < k_2 - 1$; and

$$\begin{aligned} q_{k_1} & & &= p_{k_1} - \omega \\ q_{k_1+1} &= q_{k_2-1} & &= p_{k_1+1} + 2\omega \\ q_{k_2} & & &= p_{k_2} - \omega \end{aligned}$$

when $k_1 + 1 = k_2 - 1$, where $\omega = \frac{1}{r_{m,m'}}$.

Then we have $\varphi_A = \alpha_0 q_0 + \cdots + \alpha_{R_m} q_{R_m}$. If we subtract and simplify, then

$$\varphi_A - \varphi_B = \omega(-\alpha_{k_1} + \alpha_{k_1+1} + \alpha_{k_2-1} - \alpha_{k_2});$$

which is a strictly positive quantity. In fact, the case $k_1 + 1 = k_2 - 1$ is trivial, since if $k = k_1 + 1 = k_2 - 1$ then we have

$$\omega(-\alpha_{k-1} + 2\alpha_k - \alpha_{k+1}) > 0;$$

and the remaining cases are solved by twice adding and subtracting all the intermediate coefficients

$$
\begin{aligned}
\varphi_A - \varphi_B = \omega \Big[& (-\alpha_{k_1} + 2\alpha_{k_1+1} - \alpha_{k_1+2}) \\
+ & (-\alpha_{k_1+1} + 2\alpha_{k_1+2} - \alpha_{k_1+3}) \\
& \vdots \\
+ & (-\alpha_{k_2-3} + 2\alpha_{k_2-2} - \alpha_{k_2-1}) \Big] \\
+ & (-\alpha_{k_2-2} + 2\alpha_{k_2-1} - \alpha_{k_2})
\end{aligned}
$$

where all quantities in parentheses are strictly positive. ∎

The following theoretic result establishes the unisolvency of LOP for totally $\cup$- and $\cap$-convex objective functions. Totally plane functions are not of interest because they do not vary when making interchanges, and therefore cannot be optimized (all points in the polyhedron are a solution of the LOP).

Theorem 2 *If φ is a totally $\cup$- or $\cap$-convex objective function, then the LOP has only one solution. More explicitly:*

- A: *If φ is totally $\cap$-convex, the solution of the LOP is the elemental model corresponding to a fixed composition for all (m', k')-regions in such a way that all m-regions have the same gray level or, at most, two consecutive gray levels.*
- B: *If φ is totally $\cup$-convex, the solution of the LOP is the elemental model corresponding to a fixed composition of all (m', k')-regions in such a way that all m-regions have an extreme gray level ($k = 0$ or $k = R_m$), except for, at most, one m-region in each (m', k')-region.*

Proof.

- A: According to theorem 1, every SMI causes an increase in any $\cap$-convex objective function. The optimum solution must therefore correspond to a composition which allows no SMI whatsoever. Every composition different from the one expressed must have at least two m-regions with non-consecutive gray levels, and this situation always enables us to make SMI's.

B : Similarly, every composition different from the one expressed must have at least two m-regions with non-extreme gray levels (between 1 and $R_m - 1$), and this situation always enables us to make SHI's. ■

Examples of linear models for totally $\cap$-convex and $\cup$-convex objective functions can be seen in Section III.5.

The above results indicate that the solution obtained upon optimizing a totally $\cap$-convex or $\cup$-convex objective function φ does not change when we add any totally plane function to φ. This is true despite the fact that this addition can easily move the position of the greatest coefficient of φ. Our conclusion is that the solutions of these two LOP's depend on only the shape (convexity) of the sequences of coefficients of the objective function and not on the slope of them. Both solutions will be studied in more detail in Section III.5.

III.3 RESULTS FOR PARTIALLY CONVEX OBJECTIVE FUNCTIONS

When we consider non-totally $\cap$- or $\cup$-convex objective functions we find that it is still possible to establish some results concerning the unisolvency of LOP, as shown below.

Theorem 3 Let φ be an objective function with $\cup$-convex sequences in terms $1, 2, \ldots, k_0 - 1$ and $k_0 + 1, \ldots, R_m - 1$, and with an $\cap$-convex sequence in term k_0. Let us assume that the sequence of coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{k_0 - 1}$ is strictly increasing, while the sequence $\alpha_{k_0+1}, \ldots, \alpha_{R_m}$ is strictly decreasing (see Figure 2-d). Then, the LOP is unisolvent, its solution corresponding to one of the following mutually incompatible compositions:

A : the gray level of all m-regions in the same (m', k')-region is equal to 0 or k_0, except, at most, one with an intermediate gray level ($0 < k < k_0$), or

B : the gray level of all m-regions in the same (m', k')-region is equal to k_0 or R_m, except, at most, one with an intermediate gray level ($k_0 < k < R_m$).

Proof. Cases A and B are mutually incompatible (except for the trivial case of all gray levels being equal to k_0) because the average gray level of m-regions is strictly smaller than k_0 in case A, while it is strictly greater than k_0 in case B.

Any composition other than the above two must fulfil at least one of the following conditions:

- it must have at least two m-regions with gray levels between 1 and $k_0 - 1$. In this case it is possible to make a SHI, thus increasing the value of φ.
- it must have at least two m-regions with gray levels between $k_0 + 1$ and $R_m - 1$, with same result as above.
- it must have at least one m-region with a gray level between 0 and $k_0 - 1$, and another one with a gray level between $k_0 + 1$ and R_m. This case allows us to make a SMI, thus again increasing φ. ■

An example of a linear model for this problem can be seen in Section III.5.

Curiously, a similar result for twice $\cap$-convex functions ($\cap$-convex sequences in $1;\ldots;k_0 - 1$ and $k_0 + 1;\ldots;R_m - 1$; $\cup$-convex sequence in k_0, strictly decreasing sequence in $0;\ldots;k_0$ and strictly increasing sequence in $k_0;\ldots;R_m$; see Fig. 2-e) cannot be established without specifying additional information about the coefficients α_k of the objective function. One optimum solution for this case must fulfil all of the following conditions:

- all m-regions (inside (m', k')-regions) with gray levels between 0 and k_0 must have the same gray level or, at most, two consecutive gray levels.
- the same condition for all m-regions with gray levels between k_0 and R_m.
- one (m', k')-region cannot simultaneously have an m-region with a gray level between 1 and k_0 and another with a gray level between k_0 and $R_m - 1$. (Consequently, this condition excludes the possibility of having two or more m-regions with a gray level equal to k_0).

Reasoning similar to that for the previous result leads us to conclude that these three conditions are needed; the failure of one or more makes an interchange possible (SMI for the first and second conditions, SHI for the third one), allowing an increase of φ.

However, the conditions expressed do not lead us to a single model but rather to two models, which are generally different and not mutually exclusive:

A: only gray levels appear $0, k$ and $k+1$, with $k > k_0$.

B: only gray levels appear R_m, k and $k-1$, with $k < k_0$.

For example, let us fix the values $m = 3, m' = 5, k' = 5, k_0 = 4$ ($R_m = 8$); $R_{m'} = 32$; $r_{m,m'} = 4$) so that every m'-region is composed of 4 m-regions. Solution A corresponds to the composition $0+0+0+5$, i.e. three white m-regions and one with gray level 5. Solution B corresponds to the composition $1+1+1+2$, i.e. three m-regions with gray level 1 and one with gray level 2 (no black m-region). The respective m-histograms are

$$P_A = \left\{\tfrac{3}{4}, 0, 0, 0, 0, \tfrac{1}{4}, 0, 0, 0\right\}$$
$$P_B = \left\{0, \tfrac{3}{4}, \tfrac{1}{4}, 0, 0, 0, 0, 0, 0\right\}.$$

Now let us consider the following three objective functions:

$$\varphi_1 = 10p_0 + 9p_1 + 7.0p_2 + 4p_3 + 0p_4 + 4p_5 + 7.0p_6 + 9p_7 + 10p_8$$
$$\varphi_2 = 10p_0 + 9p_1 + 7.1p_2 + 4p_3 + 0p_4 + 4p_5 + 7.1p_6 + 9p_7 + 10p_8$$
$$\varphi_3 = 10p_0 + 9p_1 + 6.9p_2 + 4p_3 + 0p_4 + 4p_5 + 6.9p_6 + 9p_7 + 10p_8.$$

Both functions verify the specified conditions. Thus we have:

$$\varphi_1(P_A) = 8.5 = \varphi_1(P_B);$$
$$\varphi_2(P_A) = 8.5 < \varphi_2(P_B) = 8.525$$
$$\varphi_3(P_A) = 8.5 > \varphi_3(P_B) = 8.475$$

For φ_1 the LOP does not have a single solution, whereas the only solution for φ_2 is different from that for φ_3.

III.4 OVERALL OPTIMIZATION

In practice, to find linear models as a solution of linear optimization problems, we must realize that one such problem appears in each column of the model. Therefore an objective function must be given for each gray level k', thus providing a set of functions to optimize $\{\varphi_0; \varphi_1; \ldots; \varphi_{R_{m'}}\}$. If we multiply the matrix of a linear model by a degenerate m'-histogram in the k'^{th} term $P_{m'} = \{0; 0; \ldots; 0; 1; 0; \ldots; 0\}$, then the resulting P_m is a copy of the k'^{th} column of the matrix. Thus, optimizing a given objective function $\varphi_{k'}$ on the k'^{th} column of the model is equivalent to optimizing it on the m-histogram corresponding to a k'-degenerate m'-histogram.

Every m'-histogram $P_{m'}$ is clearly a linear convex combination of degenerate m'-histograms. As a result of the linearity of (1), the m-histogram P_m, produced by a given $P_{m'}$, is therefore the same linear convex combination of the respective m-histograms produced by degenerate m'-histograms. In short, if we apply the model which optimizes $\{\varphi_0; \varphi_1; \ldots; \varphi_{R_{m'}}\}$ to a given input m'-histogram $P_{m'} = \{P_{m';0}; \ldots; P_{m';R_{m'}}\}$, then, since linear models preserve linear convex combinations, the resulting m-histogram is the one that optimizes the combined objective function:

$$\varphi = \sum_{k' \in \mathbf{K}_{m'}} P_{m';k'} \varphi_{k'}.$$

A natural assumption is that all m'-regions in a given image (obeying the model) are composed of m-regions, always following a common criterion, as stated by Level 2 on Section II.2. Therefore, at its own gray level k', every $\varphi_{k'}$ must convey the same characteristic of the histogram.

Let us consider the particular case in which all the objective functions in the set are identically equal. The combined objective function is also the same, and therefore P_m optimizes an objective function φ, which does not depend on $P_{m'}$. In other words, it is possible, regardless of the incoming m'-histogram, to fix an overall criterion, or objective function, for m-histograms resulting from the model. In this case all columns of the model simultaneously and independently optimize the objective function.

We shall speak generally about partial optimization when the $\varphi_{k'}$ values are not all equal. In this case, every resulting m-histogram optimizes an objective function which depends on the incoming m'-histogram. The interpretation of the combined objective function is not always straightforward.

III.5 EXAMPLES OF OPTIMIZATION

All of the examples in this section were solved for values $m = 4; m' = 6$. Thus every m'-region is composed of four m-regions.

Example 1: Totally ∩-convex objective function

This example shows a model matching the idea of low contrast in an image obtained at resolution m.

We can understand that an m-histogram indicates low contrast when its probability distribution $\{p_k\}_{k \in K_m}$ is very closely grouped. An objective function in line with this idea must outweigh the closely grouped gray levels against the sparse ones, as is true of every totally ∩-convex objective function.

The composition of each m'-region with m-regions obeys the following behavior (by gray levels):

$$
\begin{array}{lll}
0 = 0+0+0+0 & 4 = 1+1+1+1 & 8 = 2+2+2+2 \\
1 = 0+0+0+1 & 5 = 1+1+1+2 & 9 = 2+2+2+3 \\
2 = 0+0+1+1 & 6 = 1+1+2+2 & 10 = 2+2+3+3 \\
3 = 0+1+1+1 & 7 = 1+2+2+2 & 11 = 2+3+3+3 \quad \text{etc.}
\end{array}
$$

This happens in such a way that either only one gray level or two consecutive gray levels appear in every composition. We call it model B arbitrarily. Figure 3a shows a 3D-graph of the model matrix. In this figure and the following ones, depth represents the gray level k', and lines from left to right correspond to the columns of the model.

In this example the solution does not depend upon the fixed values of the coefficients of each objective function, but only on its total ∩-convexity. This fact assures that the solution for this problem is exactly the same, whether we pose an overall problem (with a maximum coefficient on middle level $R_m/2$ for all cases) or a partial problem (with a maximum coefficient on an average gray level $k' R_{m'}/R_m$ for each case). An attempt to refine the overall optimization problem by giving specific objective functions for each gray level k' did not lead us to any solution different from that of the overall case. The only determining factor in both problems has been that all coefficients in the objective functions are in ∩-convex sequences.

Example 2: Totally ∪-convex objective function.

This case is completely analogous to the previous one. The composition of every m'-region with m-regions is made in such a way that only black and white m-regions appear except, at most, one in each m'-region. Some compositions are

$$
\begin{array}{lll}
0 = 0+0+0+0 & 3 = 0+0+0+3 & 32 = 0+0+16+16 \\
1 = 0+0+0+1 & \quad\vdots & 33 = 0+1+16+16 \\
2 = 0+0+0+2 & 31 = 0+0+15+16 & \quad\text{etc.}
\end{array}
$$

We call it model Z. Figure 3b shows a 3D-graph of the matrix of the model.

All comments in the previous example hold good for this one; partial and overall problems are equivalent and the solution is the same.

Example 3: Double ∪-convex objective function.

Figure 3c shows the matrix for the solution of the overall problem, where the maximum coefficient is at the middle gray level $k_0 = R_m/2$. We call it model Y.

Figure 3d shows the matrix corresponding to the solution for the partial problem, where the maximum coefficient for each k' is at the average gray level $k_0 = \langle k \rangle = k' R_m / R_{m'}$. We call it model L.

Comments.

The graphic representation in the above figures and the mathematical computation (not given) of the matrix elements for all the examples give some information about the features of the estimated P_m histograms.

For example, the graph for model B appears to be diagonal-like. Thus, the estimated k gray level is similar (upon normalization) to that of the previous k' gray level; this model can be called “conservative” in the sense that it preserves the pattern of the input m'-histogram — for example, it provides approximately the uniform m-histogram from the uniform m'-histogram. On the other hand, since each of its k'th columns produces, at most, two consecutive k gray levels, it can also be considered being a low-contrast model.

Model Y leads to m-histograms with basically three k values: 0, R_m and $R_m/2$. This model should be appropriate for estimations in low-contrast and middle-gray-level images, since the m-histogram has an overestimated middle-gray level.

Finally, extreme gray levels are favored as much as possible in model Z, and thus should effectively model high-contrast images.

The term "contrast" as used above refers to the m-region, signifying something quite different from contrast in signal processing, which also implies changes in k' gray level.

IV APPLICATIONS IN IMAGE PROCESSING

Although the present work is mainly focused on the theoretical analysis of the probabilistic models of estimation, we shall now examine some applications of these models in image processing. The gray-value estimation (not the histogram) for the regions in the image at a certain resolution is performed on the basis of the observations of the gray values of the same image at a coarser resolution. Useful algorithms are obtained for preprocessing (filtering, smoothing, enhancement, etc.), which are adaptable to the class of images at hand. All of these multiresolution methods clearly have two features in common: 1, the image to process is not well defined, because it is either noise degraded, or not obtainable at the desired resolution in the physical system; 2, these methods involve techniques based on manipulating windows around the region to be estimated.

IV.1 METHODS AND ALGORITHMS

Now we shall examine a processing method with linear models. Let B_i be the number of observed m'-regions containing the i^{th} m-region, and let us denote their gray levels by k_{ij}; $1 \leq j \leq B_i$ (if all the possible translations of module R_m have been done, then $B_i = r_{m,m'}$ for every i). For each one of these B_i m'-regions, a new gray level k_{ij} is assigned to the i^{th} m-region, given by a central value (typically the mean, the mode or the median) of the probability distribution $\{Q(m,k|m',k_{ij})\}_{k \in K_m}$, according to the properly used model

$Q_{m;m'}$. This computation constitutes the first stage in the process. The second stage consists of taking again a central value, this time for the set $\{k_{ij}\}_{j=1}^{B_i}$ of central values obtained in the first stage from the observed m'-regions. In this way up to nine cross-strategy options for computation should have been obtained. We can use an order statistic, apart from the median, for an even greater variety of strategies in order to achieve a given goal.

The above procedures can generally be viewed as nonlinear filters. The particular application to a given image may be preprocessing, restoration, gray-level smoothing, edge sharpening, and others. When filtering out noise from images, the particular strategy mean+median resembles the well-known median filters, since both use window and estimation by the median. The general procedure offers many options and parameters, depending on the operation: mainly the two-stage strategy, the pair $(m; m')$ of resolutions, and the model can be set up. The particular properties of these new filters are currently investigated by the present authors.

The filter operation described above can be applied for two practical situations which differ from each other in the way they lack reliable information. In the first situation, the image can be obtained just at resolution m', while it is required at resolution m. Thus, the gray level of the m-regions should be estimated. The idea is to supply the lack of information with an extra number of observations at resolution m'. The procedure must result in an m-image from a set of m'-images coming from the original one by different, slightly-scrolled views. This method may be called superresolution. In the second practical situation, the lack of information comes from the presence of noise in the image, which is obtained at the desired resolution. The filter operation now maps an image into another one at the same resolution with improved quality.

IV.2 EXAMPLES

The first example concerns the superresolution operation. The algorithm works on 16 views of the picture at resolution $m' = 12$, obtained by shifting the observation system by $R_{m'}$-shaped jumps in each step.

Figure 4 shows a computer-simulated example. The "views" are obtained by adding the gray levels of the m-regions within the (simulated) m'-windows of the image. A single sample of this set of views is presented in part a. Parts b;c show the results after 1 and 2 processing iterations, respectively.

Median+median strategy and model B have been used. The second iteration does not improve the result in this example, while the blurring effect is increased.

It is worthy of note that the above algorithm can also work by randomly shifting the observation window, instead of regularly, keeping the jump size in the average. The result can be somewhat worse, although it may be balanced with a greater amount of views.

The second example shows some cases of filtering a synthetic image, corrupted by impulsive noise in order to display the ability of the model-based algorithms to remove it. The filter operation is illustrated in Figure 5. Parts $a; b$ show the original and noisy (15% of impulsive noise) images. Parts $c; d; e$ presents the results obtained with models $B; Y; Z$. Parts $f; g; h$ show the results after three passes of the filtering algorithm with the same models as before. Finally, parts $i; j; k$ show a different original, noisy and filtered image by using model B. Median+median strategy has been used throughout.

The third example is another case of removing impulsive noise. The filter operation is shown in Figure 6, in which the original, the noisy source to process and the results with the strategies and models as indicated, are presented in parts $a; b; c; d$. Large differences can be appreciated, depending on the strategy used; only the two strategies with the best results have been carried out. On the other hand, the model strongly determines the performance; the given picture calls for model Z because of its high contrast.

Some remarks about applications should be made about the described examples.

1. The results are strongly depending upon the used model, as expected from their theoretical construction. While model B (conservative) preserves the gray levels, model Z increases the contrast and model Y biases toward the middle gray.

2. Model Z performs the best in removing the impulsive noise in images with high contrast, since edges between similar light (or dark) levels are lost.

3. Although median+median strategy is the most useful, mode+mode has shown to achieve better results for filtering a binary yet strongly corrupted image (Figure 6).

4. When using model B, a few iterations filter out the impulsive noise further, but increase the blurring effect in the superresolution operation.

CONCLUSIONS

The interdependence of gray-level histograms in a multiresolution framework has been studied for digital images, in which gray-value additivity is involved when composing regions. Probabilistic linear models for this interdependence have been developed, which are applied to estimate histograms. By defining the elemental models and taking them as a basis, the space of the consistent models is generated. Linear-optimization methods are then used for designing models, according to the prior information available about the proper class of images.

Applications of these models in image processing are presented in the general framework of nonlinear filtering, showing the high versatility of the procedures. The two-stage, model-based, order-statistic filter has received particular attention, especially the median type. Some examples are shown.

Finally, this analysis can be extended to signals other than 2-D images, provided that the additivity of the signal values is maintained throughout the multiresolution relationships. Applications could also be extended to many other fields, such as texture analysis, segmentation, feature extraction and statistical pattern recognition.

Acknowledgements

We thank Juan Francisco Gómez-Lopera for collaborating in the preparation of the experiments.

References

[1] CASTLEMAN, K.R. Digital Image Processing, Prentice Hall, New Jersey 1979.

[2] CIBULSKIS, J.M., DYER, C. An analysis of node linking in overlapped pyramids. IEEE Trans. on Sistems, Man and Cybernetics, SMC-14, 3, 424-436 (1984).

[3] CLIPPINGDALE, S.C., WILSON, R.G. Least-squares image estimation on a multiresolution pyramid, Proc. ICASSP '89, 1409-1412.

[4] GARBOOJ, M., HAAVISTO, P., NEUVO, Y. Recent advances in median filtering. Communication, Control and Signal Processing, (Proc. of "Bilcon'90). E. Arikan (Ed.), Elsevier Science Publishers B.V., Amsterdam, 1990, 1378-1385.

[5] GONZALEZ, R.C. Image Enhancement and Restoration, Handbook of Pattern Recognition Image Processing, Academic Press, San Diego (CA), 1986.

[6] GROSKY, W.I., JAIN, R. A pyramid-based approach to segmentation applied to region matching. IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-8, 5, 639-650 (1984).

[7] HONG, T.H., SHNEIER, M. Extracting compact object using linked pyramids. IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-6, 2, 229-236 (1984).

[8] JAIN, ARIL K. Fundamentals of digital image processing. Prentice-Hall Int. Ed. New Jersey 1989.

[9] LINDEBERG, T. Scale-space for discrete images, Contributions to the 6th Scandinavian Conference on Image Analysis, Royal Institute of Technology, Stockholm, Sweden. Report TRITA-NA-P8905 CVAP 61.

[10] LINDEBERG, T. Scale-space for discrete signals. IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-12, 3, 234-254 (1984).

[11] MARTÍNEZ-AROZA, J. Contribución al análisis y estimación en imágenes digitales mediante la teoría de la información. Ph.D. Thesis. October 1990. Publicaciones de la Universidad de Granada.

[12] MARTÍNEZ-AROZA, J., QUESADA-MOLINA, J.J., ROMÁN-ROLDÁN, R. A characterization of images through entropy-resolution diagrams. Signal Processing V. Theories and Applications. L. Torres

& al. (Ed.). Elsevier Science Publisher B.V., 1990, 1055-1058. (Proceedings of EUSIPCO-90, Barcelona, Sept. 1990). ISBN 0-444-88636-2.

[13] MARTÍNEZ-AROZA, J., QUESADA-MOLINA, J.J., ROMÁN-ROLDÁN, R. Estimation of images via conditional probabilities. Communication, Control, and Signal Processing, E. Arikan (Ed.), Elsevier Science Publishers B.V., Amsterdam 1990, 1378-1385. (Proceedings of the 1990 Bilkent Int. Conf. on "New Trends in Comm., Cont., and Sig. Pro." Bilcon'90, Ankara, July 1990). ISBN 0-444-88762-8.

[14] MARTÍNEZ-AROZA, J., QUESADA-MOLINA, J.J., ROMÁN-ROLDÁN, R. Image estimation through probabilistic methods. Proceedings of the IV Symposium Nacional de Reconocimiento de Formas y Análisis de Imágenes. Granada, Sept. 1990. 324-326.

[15] MARTÍNEZ-AROZA, J., QUESADA-MOLINA, J.J., ROMÁN-ROLDÁN, R. Image processing methods based on linear models of interresolution dependence. Proceedings of the China 1991 International Conference on Circuits and Systems. IEEE 1991. Pg. 552-555. ISBN 0-7803-0150-2.

[16] PITAS, I., VENETSANOPOULOS, A.N. Nonlinear Digital Filters. Principles and Applications. Kluwer Academic Publishers. The Netherlands, 1990. ISBN 0-7923-9049-0.

[18] ROMÁN-ROLDÁN, R., QUESADA-MOLINA, J.J., MARTÍNEZ-AROZA, J. Multiresolution estimation of gray-level histograms. Pattern recognition and image processing in physics, R.A. Vaughan (Ed.). Adam Hilger, Bristol (UK) 1991, 343-348. (Proceedings of the NATO ASI, Dundee (UK), August 1990). ISBN 0-7503-0121-X.

[19] ROMÁN-ROLDÁN, R., QUESADA-MOLINA, J.J., MARTÍNEZ-AROZA, J. Multiresolution-information analysis for images. Signal Processing 24 (July 1991), 77-91 (Elsevier).

[20] SPANN, M., WILSON, R. A quad-tree approach to image segmentation which combines statistical and spatial information. Pattern Recognition 18, 3/4, 257-269 (1985).

[21] TAKACS, L. Combinatorics, Handbook of Statistics 4, P.R. Krishnaiah & P.K. Sen (Eds.), North-Holland Science Publishers, Amsterdam, 1984, 123-173.

[22] WONG, A.K.C., VOGEL, M.A. Resolution-dependent information measures for image analysis, Trans. IEEE on Systems, Man and Cybernetics SMC-7, 1, Jan. 1977.

FOOTNOTES (first page)

José Martínez-Aroza is with the Department of Applied Mathematics, University of Granada, 18071 Granada, Spain. Tel. 58 24 31 92, Fax 58 24 32 86, e-mail **JAMARTINEZ@ugr.es**.

Ramón Román-Roldán is with the Department of Applied Physics, University of Granada, 18071 Granada, Spain. Tel. 58 24 32 22, Fax 58 24 32 14, e-mail **RROMAN@ugr.es**; IEEE Log Number 1009349.

This work is supported in part by the C.I.C.Y.T. under Research Grant TIC 91-0646.

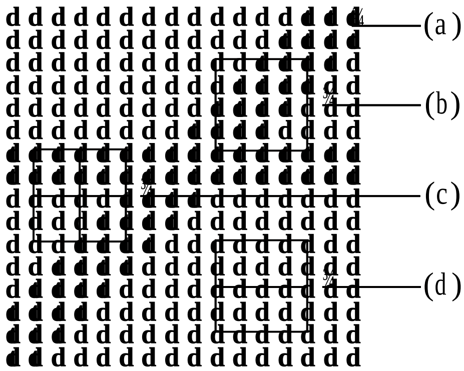

Figure 1: (a) a black pixel; (b) a 4-region in which the gray level is 12, i.e., a (4,12)-region); (c) a 4-region partitioned into four 2-regions; (d) a (4,0)-region partitioned into two (3,0)-regions.

Figure 2: Sequence of coefficients α for an objective function φ: (a), totally plane; (b), totally $\cup$-convex; (c), totally $\cap$-convex; (d), double $\cup$-convex; (e), double $\cap$-convex.

Figure 3: 3-D representation of the matrix elements for model: (a), conservative B; (b), high-contrast Z; (c), overall low-contrast Y; (d), partial low-contrast L; (e), hypergeometric H.

Figure 4: An example of superresolution from $m' = 12$ to $m = 8$, with model B and median+median strategy. (a), a sample among 16 of the original image; (b), result after 1 iteration; (c), result after 2 iterations.

Figure 5: Examples of filtering out impulsive noise from $m' = 8$ to $m = 4$. (a), original picture (for comparison only); (b), damaged picture with 15% impulsive noise (the source); (c;d;e), results after 1 iteration with strategy median+median, using models B,Y,Z, respectively; (f;g;h), same as (c;d;e), after 3 iterations; (i), original picture (for comparison only); (j), damaged picture with 15% impulsive noise (the source); (k), result after 1 iteration with strategy median+median, using model B.

Figure 6: An example of filtering out impulsive noise from $m' = 8$ to $m = 2$. (a), original picture; (b), damaged picture with 90% impulsive noise (the source). (c), result by using model Z, mode+mode strategy; (d), result by using model Z, median+median strategy.