

AUTOMATIC BUILDING DETECTION FROM HIGH RESOLUTION IMAGES BASED ON MULTIPLE FEATURES

Zhen Ge. Qiu, *IEEE Member*^{a,*}, Shi Tao. Zhang^b, Chun Ling. Zhang^c, Jin Yun. Fang^a

^a The Institute of Computing Technology of Chinese Academy Of Sciences, No.6 South Ke Xue Yuan Road, Beijing, China - qiuzhenge@ict.ac.cn

^b University of Science and Technology Beijing, Beijing, China - changong-ok@263.net

^c Surveying & Mapping Bureau Of He Nan Province ,China

Commission II, IC WG II/IV

KEY WORDS: Edge detection; Earth Mover's Distance; Curvature; Fuzzy density; Hough transform; Building Detection;

ABSTRACT:

Automatic recognition and reconstruction of buildings from aerial and space images is of great practical interest for many of applications such as cartography and photo-interpretation. Building detection is the first and very difficult step in building recognition and reconstruction. It is to find buildings and separating them from the background in the presence of distractions caused by other features such as surface markings, vegetation, shadows and highlights. This is an instance of the well-known figure-ground problem. The goal of automatic building detection in this paper is to roughly delineate the rooftop of the buildings that will be verified during the recognition and reconstruction phase. The rooftop detection algorithm proposed here is based on multiple features and proceeds in two steps: first, low-level feature extraction; second, rooftop identification. In this paper we focus on rectangle building roof recognition. In this case, the boundaries of their rooftop are straight lines. One of the obvious facts is that most build roofs are built of materials of limit categories, so their image colors and textures are of limited categories. Low-level features used here are straight-line segments, image colors and image textures. In local edge detection, the vital phase of low-level feature extraction, we introduced a novel edge detection algorithm based on EMD (The Earth Mover's Distance), which works better than traditional ones. A general curvature concept was used in the measuring of image textures, which is invariant to rotation. And at last with the mathematics tools of Hough Transformation and fuzzy density function we made the last decision to determine "it is building or not". Experiments were carried out in the Quick-bird images of Beijing, China. We were able to achieve a right detection rate larger than 75% for those buildings that are not occluded severely.

1. INTRODUCTION

The measure of locating areas of city changes through remotely sensed data is an important tool for many applications including resource management, urban planning, and environmental impact analysis, especially in the fast developing areas in China. In fact, in China, monitoring land cover changes is currently one priority of city land management bureau's works in the coming years. To efficiently monitor land cover changes, we need an efficient mean of determining where is a new building. Till now, although many experts of diverse background have struggled for many years to create such kind of tools, there is still no satisfied one. We have to identify buildings in images with tedious manual editing. In this paper, we start a new try to develop a tool of automatic recognition of buildings from aerial and space images without or with minimal operator interaction. As to previous works in this field, Prof.Grun, etc, made a excellent review of detection of buildings from aerial space images(Grun *et al*, 1995; Robert W. Carroll), some more work is described in (Fua, 1996; Henricsson *et al*, 1996; Weidner, 1996). In the last several years, more and more new kinds of inputs, such as stereo images and range images are used. In this paper, we still focus on the use of single-images. With little of direct 3-D information, making use of single-images is more difficult, but it is attractive due to the ease with which they can be obtained. Further more, many of the processes involved in

single-image analysis are also required for multiple-image analysis.

Our algorithm is restricted to recognize the rectangular rooftop of the buildings in single-images, especially in color-images. The rooftop detection algorithm proposed here is based on multiple features and proceeds in two steps: first, low-level feature extraction; second, rooftop identification. Low-level features are straight-line segments and image texture. As local edge detection is vital here; in colour images, we detect local edge points with a novel edge detection algorithm based on EMD (The Earth Mover's Distance), which works better than traditional ones. After detecting local edges, straight-line segments are found with Hough transformation. The probability density of image curvature is used as the feature of texture in an image patch, which is invariant to rotation. In the rooftop identification step, we first find potential rectangles by using of the parameters space that is got in Hough transformation. Based on the ratio of width and height of one rectangle, the size of the rectangle, the properties of the probability density of local image's curvature in the rectangle and the average color of local image in the rectangle, we compute fuzzy integrals based on the evidences gathered to confirm whether the area belongs to a rooftop. Our novel strategy improves the efficiency of human work in building recognition with few interactions.

The paper is organized as follows. First we give a relative detail presentation of the novel edge detection algorithm based on EMD (The Earth Mover's Distance) in section 2. Second, in

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

section 3, we picture the main idea of how we find Straight lines and Rectangles with Hough transform. Third, in section 4, the formulas to calculate Curvature of color-images are outlined. Last, in section 5, we describe the way to recognize rooftop by using of fuzzy density.

2. NOVEL EDGE DETECTION ALGORITHM OF COLOR-IMAGES BASED ON EMD

As the edge detection is one of the bases of our buildings detection algorithm, we made much more efforts on it. After many tests and comparisons, we introduce Earth Mover's Distance to the buildings detection algorithm in remote sensing images.

With few exceptions, the fundamental assumption of all step edge detectors is that the regions on either side of an edge are constant in colours or intensity. Much effort has gone into making them robust to noise, but the noise is assumed to have statistically simple properties.

Convolution masks are ideal for realizing this assumption because the sign of the weight at a pixel tells us what side of the edge it is hypothesized to be on. We can think of a convolution as finding the weighted mean of each side and then computing the distance between the two means (Eric N. Mortensen, 2001; Mark A. Ruzon, 1999).

While this assumption holds well enough for many applications, it does not hold in all cases. For instance, as scale increases, it is more likely that the weighted mean of each side will not be meaningful because an operator will include image features unrelated to the edge. This observation is even truer of color images. When only intensities are involved, the average over a large window is still perceptually meaningful because intensities are totally ordered. In color images, there is no such ordering, so the "mean color" of a large window may have little perceptual similarity to any of the colors in it.

To overcome these barriers, we introduced a novel edge detection algorithm based on Earth Mover's Distance to color remote sensing images. As reported recently, Earth Mover's Distance works well in edge detection of images used in other fields. This novel algorithm proposes an edge model that assumes that the two distributions of pixel values on either side of an edge are different. Distributions allow for more control over how each region is represented and how the distance between two regions is computed than can be achieved by using only the mean value.

Besides the above, this edge model has other advantages. The first is a lack of false negatives compared to other models false negatives result from a failure to "take into account all possible intensity variations that might accompany a step edge in practice". Since we use distributions, almost all of these variations are modelled implicitly. In homogeneities can be uncorrelated (due to noise) or correlated (due to texture) without affecting performance. The second benefit is that using distributions creates a unifying framework for edge detection in binary, grey-scale, color, or multi-spectral images, so long as a meaningful ground distance is defined. So our algorithm is a generalist in edge detection.

2.1 A Perceptual Ground Distance

One of the key concepts in our novel edge detection algorithm is the distance between two color signatures. Before defining the distance between two color signatures, we must first define the ground distance between two colors. Because this distance should conform well to human perceptual distance as measured by psychophysicists, we use the CIE-Lab color space, in which

small Euclidean distances are perceptually accurate. If two colors are separated by a long distance, however, that distance is no longer quantitatively meaningful; the most we can say about the colors is that they are different. Using the Euclidean distance by itself presumes that an edge with a contrast of 80 units is twice as salient as an edge with a contrast of 40 units, which need not be true. We desire a distance measure that approaches but does not exceed 1 once the colors are far enough apart. There are many functions that satisfy this criterion, and we have chosen (Mark A. Ruzon, 1999; G. Wyszecki., 1982)

$$d(i, j) = 1 - \exp(-E_{ij}/r) \quad (1)$$

where E_{ij} is the Euclidean distance between color i and color j .

r is a constant that determines the steepness of our function. We have empirically chosen $r = 16.0$ for our experiments.

This information has more psychophysical background and the output based on this is more near to what of human sight.

2.2 The Earth Mover's Distance

We introduce a new distance between two signatures that we call the Earth Mover's Distance (EMD). This reflects the minimal cost that must be paid to transform one signature into the other. EMD is a general method for matching multidimensional distributions. The main idea of EMD is presented below (S. Cohen and L. Guibas, 1999):

Consider a set of points $L = \{l_1, \dots, l_n\}$ with a distance function $d(i, j)$ (assumed to be a metric). A distribution $P(L)$ on L is a collection of non-negative weights (p_1, \dots, p_n) for points in X such that $\sum p_i = 1$. The distance between two distributions $P(L)$ and $Q(L)$ is defined to be the optimal cost of the following minimum transportation problem:

$$\begin{aligned} \min \sum_{i,j} f_{i,j} \cdot d(i, j) \\ \forall i \quad \sum_j f_{i,j} &= p_i \\ \forall j \quad \sum_i f_{i,j} &= p_j \\ \forall i, j \quad f_{i,j} &\geq 0 \end{aligned} \quad (2)$$

Above we define a somewhat restricted form of the Earth Mover Distance. The general definition does not assume that the sum of the weights is identical for distributions $P(L)$ and $Q(L)$. This is useful for example in matching a small image to a portion of a larger image.

2.3 Computing Edge Information

We detect local edge in a circle neighbour area. As depicted in Fig.1 by dividing the circle in half, we have created two color signatures of equal mass, which we denote S_1 and S_2 . Finding the distance between them can be seen as an instance of the transportation problem, in which we wish to find the minimum amount of physical work needed to move the masses of S_1 into correspondence with those of S_2 in color space. The Earth Mover's Distance (EMD) is based on a solution to

this problem. It has been proven to be more robust than other distance measures when comparing the color signatures of entire images because it avoids many quantization and discretization influences.

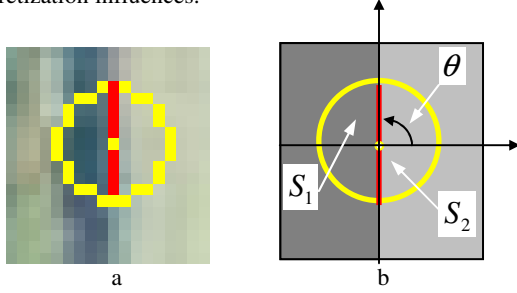


Figure 1. Detect local edge in a circle neighbour area

We can now summarize the algorithm. For every θ (the edge's orientation at the centre point), there is a division of S_1 and S_2 , so the resulting of EMD can be represented as a function $f(\theta)$, $0 < \theta < 180^\circ$. $f(\theta)$ is one period of a triangular wave. We define the orientation at the centre point to be $\bar{\theta} = \arg \max_{\theta} f(\theta)$, and the strength to be $f(\bar{\theta})$. While the importance of the maximum is intuitive, the minimum is equally important. Regardless of strength, the minimum may still be zero if there is an orientation produces two equal color signatures. The minimum measures the photometric symmetry of the data; when it is high, our edge model is violated. For this reason, the value $\min_{\theta} f(\theta)$ is called the abnormality. One cause of high abnormality is the existence of a junction (Mark A. Ruzon, 1999).

3. HOUGH TRANSFORM FOR STRAIGHT LINES AND RECTANGLES DETECTION

One of the key work phases here is to find parallel lines. In our application; first we used the standard Hough transform for lines, and then find parallel lines with the lines orientation angle recorded in the parameter space of HT. Since the Hough transform (Mark C.K. Yang, Jong-Sen Lee, 1997) has been successful in detecting lines, circles, and parabolas. We can represent a line parametrically as follows:

$$\rho = x \cos \theta + y \sin \theta \quad (3)$$

For every non-zero pixel within the region, we can consider its coordinate (x, y) in the parameter space (ρ, θ) called the Hough space. An array H is generated for representing the Hough space, which is divided into regular lattice regions referred to as cells. If a cell is at the intersection of many curves in the Hough space it is said to have a large number of votes (for a true line). From Eq. (3) it is not difficult to see that the cell indicates, potentially, there is a line in the image domain. The size or the resolution of the accumulator array H plays an important role in the process and must be carefully designed. For the buildings can head in any direction, thus the range of H is

$$\begin{aligned} 0 < \theta < \pi \\ -s < \rho < s \end{aligned} \quad (4)$$

where s is the diagonal of the sub-image containing region of interesting.

After getting the accumulator array H , count the number of point in each cell. Here we have empirically chosen a minimum threshold of 10 for our experiment, which means that any line which has lesser than 10 points will be ignored. We do this by simply setting the cell value to zero if its original value smaller than 10.

Next work is a bit complex, where we will find potential rectangles. First we defined two thresholds: the minimum of the length and width of rectangle $\min \rho$ and the maximum of the length and width of rectangle $\max \rho$. Second, in each line of the array H (of the same θ), find cell pairs meet the condition that the ρ distance (for example showed in Fig 2: $\Delta\rho1, \Delta\rho2$) between them is larger than $\min \rho$ and smaller than $\max \rho$. The cell pairs are the parallel lines we defined. Then we find couples of cell pairs' in different lines, which meet the condition that the θ distance ($\Delta\theta$, as depicted in Fig.2) between the lines is larger than 87° and smaller than 93° (which is given naively in this experiment). Last we find the potential rectangle with these couples of cell pairs'. After detecting a potential rectangle, we extract some features (height, width, size) of it.

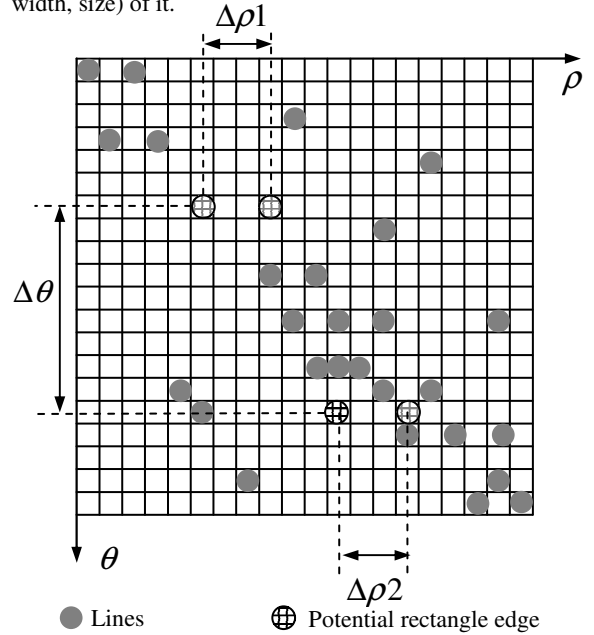


Figure 2. Find potential rectangle in parameter space

4. USING CURVATURE AS THE IMAGE TEXTURE MEASURE

Curvature is powerful concepts used in describing the instinct feature of surface and it is invariant to rotation. Here we use it to measure the texture of a patch in color-images. The curvature and relative concepts we used here are defined as below

First, give some related differential geometry concepts (Monge Patch, Mathworld, USA):

Definition of Patch: A patch (also called a local surface) is a differentiable mapping $x: U \rightarrow R^n$, where U is an open subset of R^2 . More generally, if A is any subset of R^2 , then a map $x: A \rightarrow R^n$ is a patch provided that x can be extended to a differentiable map from U into R^n , where U is

an open set containing A . Here, $x(U)$ (or more generally, $x(A)$) is called the map trace of x .

Definition of Monge Patch: A Monge patch is a patch $x: U \rightarrow \mathbb{R}^3$ of the form

$$x(u, v) = (u, v, h(u, v)) \quad (5)$$

where U is an open set in \mathbb{R}^2 and $h: U \rightarrow \mathbb{R}$ is a differentiable function.

Definition of Gaussian curvature K and Mean curvature H :

For a Monge patch, the Gaussian curvature K and Mean curvature H are

$$K = \frac{h_{uu}h_{vv} - h_{uv}^2}{(1 - h_u^2 + h_v^2)^2} \quad (6)$$

$$H = \frac{(1 + h_v^2)h_{uu} - 2h_uh_vh_{uv} + (1 + h_u^2)h_{vv}}{2(1 + h_u^2 + h_v^2)^{3/2}}$$

where $h_u, h_v, h_{uu}, h_{uv}, h_{vv}$ are Partial derivatives of $h(u, v)$.

Here, then, we generalize the h to a map from an open set in \mathbb{R}^2 to \mathbb{R}^3 ; $h^{color}: U \rightarrow \mathbb{R}^3$, and think that any arbitrary point in \mathbb{R}^3 represents a color in CIE-Lab color-space. As defined in section 2, the distance between two colors can be measured by Eq. (1). Based on the definition of differentiation in \mathbb{R}^n , we defined partial derivatives of $h^{color}(u, v)$ as derivatives of the function h^{color} when all but the variables of interest are held fixed during the differentiation (Wilhelm Klingenberg, 1978).

Definition of Mean Curvature Of Color-image:

$$h_u^{color} = \frac{\partial h^{color}}{\partial u} \quad (7)$$

$$= \lim_{\Delta u \rightarrow 0} \frac{d(color(u + \Delta u, v) - color(u, v))}{\Delta u}$$

Where $color(u + \Delta u, v)$ and $color(u, v)$ are the i, j in Eq. (1) respectively. Assume h^{color} is a "nice" two-dimensional function, so $h_u, h_v, h_{uu}, h_{uv}, h_{vv}$ exist. With these we can calculate the Mean curvature H as Mean Curvature Of Color-image.

At last, in our experiment, we used the Mean curvature H as the property of the color-image texture. After finishing the work of section 2 and 3, we got many potential rectangles; we think these potential rectangles as the patches in which we calculate the curvatures and histogram of curvatures, the latter represents the probability density of local image's curvature.

5. RECOGNITION USING FUZZY INTEGRALS

After the complex pre-proceedings, we got such datum: No.1 is the rectangles, No.2 is the histogram of curvatures. As to rectangles, we use these features of them: height, width, size; as to the histogram of curvatures, the features that we employ here are: high mean absolute deviation and low mean absolute deviation. High mean absolute deviation is the mean absolute

deviation calculated with the whole histogram of curvatures. Low mean absolute deviation is the minimum one calculated by

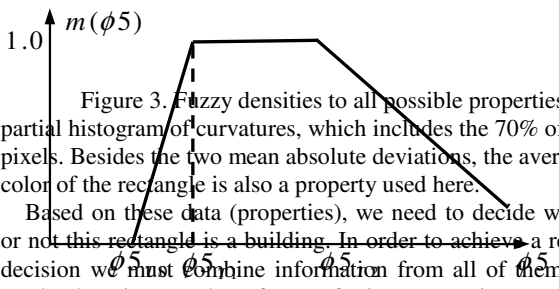
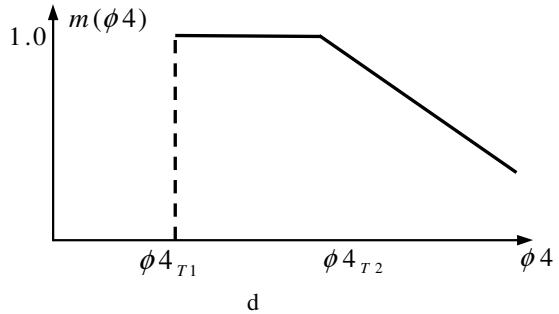
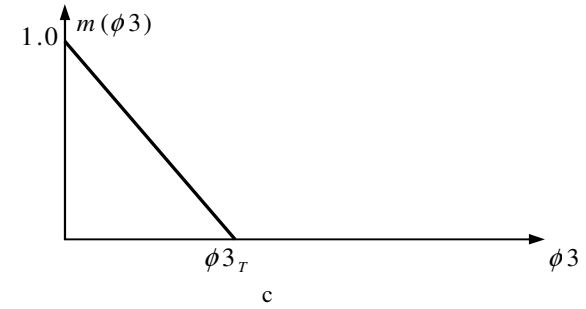
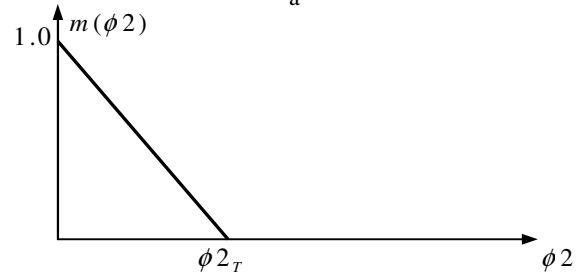
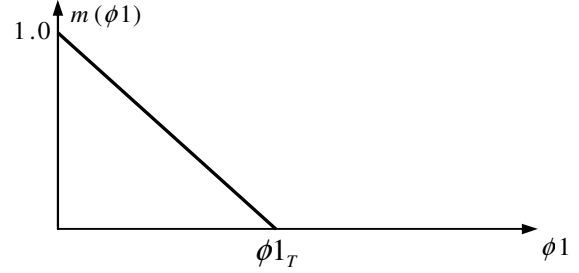


Figure 3. Fuzzy densities to all possible properties partial histogram of curvatures, which includes the 70% of total pixels. Besides the two mean absolute deviations, the average color of the rectangle is also a property used here.

Based on these data (properties), we need to decide whether or not this rectangle is a building. In order to achieve a reliable decision we must combine information from all of them. This can be done in a number of ways; for instance, using a Bayesian approach (Z.Q. Liu, 1997), Dumpsters-Shafer (G. Shafer, 1976), or fuzzy logic (J.C. Bezdek, 1999.). One technique that has enjoyed success in other vision applications is the use of fuzzy measures. In our approach, we follow this method.

We first treat each property (information) as fuzzy variables and assign fuzzy density to all possible values. Fig.3 (a) is for property $\phi 1$, which is the minimum distance from the average color of the region to the prior-colors selected empirically,

which most of building roofs would probably present. There is a reasonable hypothesis that most of build roofs are created with materials of limit categories, so they have limit kinds of color and texture. Fig.3 (b) and Fig.3 (c) are for properties ϕ_2 and ϕ_3 , which are the high mean absolute deviation and low mean absolute deviation of the histogram of curvatures. Fig.3 (d) is for property ϕ_4 , which is the ratio of width over height of the region. We assert the two thresholds are 0.19 and 0.45 respectively in our test. Fig.3 (e) is for property ϕ_5 , which is the size of the rectangle. Most of the properties' thresholds except ϕ_4 are adapted to the images to be processed; according to our experience, they are relatively common to the same batch of images.

As a simple fuzzy integral, last decision is made by naively computing the average of the fuzzy density to all possible properties, and assert a threshold to determined "it is a building" or "it is not a building".



Figure 4. One of tests in the images of Beijing, China

6. CONCLUSIONS

As shown in Fig.4, we made some experiments on some color high-resolution images; it is one of the parts of an image of Beijing area. The red line rectangles are the parts identified as buildings by our algorithm. Verified in the available image samples, we achieve a right identify ratio lager than 75%.

References from Journals:

Eric N. Mortensen, 2001. A Confidence Measure for Boundary Detection and Object Selection. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Fua P.,1996. Model-Based Optimization: Accurate and Consistent Site Modeling. *Proceedings of the 18th SPRS Congress*, Comm. III, WG 2, Vienna, Austria, pp. 222-233.

Henricsson O., F. Bignone, W. Willuhn, F. Ade, O. Kubler, E. Baltsavias, S. Mason, A. Grun, 1996. Project AMOBE: Strategies, Current Status and Future Work. *Proceedings of the*

18th SPRS Congress, Comm. III, WG 2, Vienna, Austria, pp. 321-330.

Mark C.K. Yang, Jong-Sen Lee, 1997. Hough Transform Modified by Line Connectivity and Line Thickness. *Ieee Transactions On Pattern Analysis And Machine Intelligence*, Vol. 19, No. 8, pp. 905-910

S. Cohen and L. Guibas, 1999. The Earth Mover's Distance under Transformation Sets. *Proc.7th IEEE Intel.Conf. Computer Vision*.

Weidner U.,1996. An Approach to Building Extraction from Digital Surface Models. *Proceedings of the 18th SPRS Congress*, Comm. III, WG 2, Vienna, Austria, pp. 924-929.

Z.Q. Liu, 1997. Bayesian paradigms in image processing. *Int. J. Pattern Recognition Artif. Intell.* 11 (1),pp.3-34.

References from Books:

G. Shafer, 1976. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.

G.Wyszecki., 1982. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley and Sons, New York, NY.

J.C. Bezdek, J.M. Keller, R. Krishnamupram, S.K. Pal, 1999. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Kluwer Academic Publishers, Norwell, MA.

Wilhelm Klingenberg, 1978. *A Course In Differential Geometry*. Springer-Verlag.pp.3.

References from Other Literature:

Grun A., O. Kubler, P. Agouris, 1995. Automatic Extraction of Man-Made Objects from Aerial and Space Images, *Virkhauser Verlag*, Basel, pp. 199-210.

Mark A. Ruzon, 1999. Color Edge Detection with the Compass Operator. Computer Science Department Stanford University Stanford, CA 94305.

Robert W. Carroll. Report of Detecting Building Changes Through Imagery And Automatic Feature Processing. Hitachi Software Global Technology, Ltd.10355 Westmoor Drive, Suite 250 Westminster, Colorado 80021 303-466-9255, Rcarroll@HSGT.com

References from websites:

Monge Patch, Mathworld, USA
<http://mathworld.wolfram.com/MongePatch.html> (accessed 22 Apr .2004)