# Chapter 7
# Classification of Building Images in Video Sequences


## ABSTRACT

*A technique for detection of building images in real-world video sequences is presented. The proposed technique uses information extracted from video features to improve precision in classification results. It combines fuzzy rule-based classification with a method for changing region detection in outdoor environments, which is invariant to extreme illumination changes and severe weather conditions. It has been tested on sequences under various lighting conditions. Satisfactory and promising results have been achieved.*


## 1. INTRODUCTION

The problem of edge-based building image classification in real-world video sequences is addressed in this section. The presented technique exploits information in video primitives to focus classification on features extracted from pixels belonging to static regions.

The introduced technique is based on two simple observations: In static regions the possibility of finding features that match the pattern of "building" is higher because buildings are rigid and static objects; and misclassification can be reduced removing pixels belonging to changing regions. These regions can contain objects semantically different to buildings but with an edge distribution highly similar, i.e. high frequency of vertical and horizontal edges.

Existing approaches for classification of building images use a Bayesian framework to exploit image features by perceptual grouping (Iqbal, 1999), binary Bayesian hierarchical classifiers (Vailaya, 2001), or perform building semantic extraction using support vector machines (Wang, 2002).

## 2. EDGE-BASED BUILDING IMAGE CLASSIFICATION

Edge-based building image classification can be found in (Iqbal, 1999), (Vailaya, 2001), (Wang, 2002).

Let x be an image or video frame, $f = \{f(1), ..., f(n)\}$ be feature sets where f is a function of the image x, $Et = (e_1t , ..., e_5t )$ be a pattern extracted from a feature vector $f^{(t)}$, and $Y = \{B, NB\}$ be a class set (B:Building, NB:Non Building). Edge-based building image classification uses a function:

DOI: 10.4018/978-1-4666-4896-8.ch007

$$g(x): E^t \rightarrow Y \quad (1)$$

where $E^t$ is extracted from the description of x given by the MPEG-7 edge histogram descriptor.

This descriptor uses an 80-bin histogram to represent the local distribution of directional (vertical, horizontal, diagonal 45°, and diagonal 135°) and non- directional (isotropic) edges.

g(x) is approximated with a set $R = \{R_1, \ldots, R_C\}$ of if-then fuzzy rules which has the following structure:

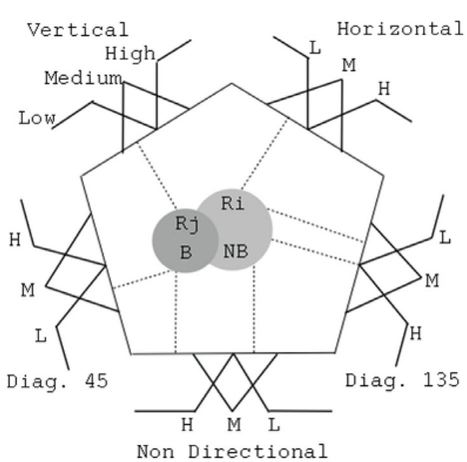
$$^{W}R_k\text{: If } e_i^t \text{ is } A_i^k \text{ then Y is } y_j \quad (2)$$

where $^{w}R_k$ is a weighted rule, $w \in [0, 1]$ is the weight, $e_i^t$ a specific type of edge, and $A^k$ is a linguistic label used to transform values from a continuous to a discrete domain.

The fuzzy model has five input variables because of the semantics of edge histogram descriptor which uses five types of edges. Besides, a group of three fuzzy sets is associated with each variable. Thus, rule base is a multi-input single-output space as is depicted in Figure 1.

$R_i$ and $R_j$ are if-then rules associated with classes NB and B, respectively. H, M, L stand for High, Low, Medium, and Low.

*Figure 1. Fuzzy model for edge histogram descriptor*




After several experiments, it was identified that most of the misclassifications of building images in real-world videos were due to objects with a similar edge distribution of buildings but semantically different. Most of them were moving non-rigid objects. Therefore, the technique was improved integrating a pre-processing method to detect changing regions.

## 3. CHANGING REGION DETECTION

A shading model method for moving object detection is applied in order to perform detection of changing regions. This version is invariant to extreme illumination changes at pixel level.

It uses the ratio of pixel intensities within a sliding mask $A_i$, i = 1 . . . N, which performs scanning of a window W in each frame within a video sequence I. W has fixed size and position, the pixel variance $^{I}\sigma^2$ is defined as described in the Equation (2) in Chapter 2.

$K_i$ is an adaptive coefficient to overcome falsely assignment of pixel to changing region for fast and large illumination changes. $K_i$ is defined according to the equation (3) in Chapter 2.

In this way, changing regions are located and used to obtain a binary image $I_b$ that contains white and black pixels to represent changing and static regions, respectively.

As an example, frames 50850 and 27172 showed in Figures 2(a) and (c), respectively, are classified as "building image". Looking at Figures 2(b) and (d), edge histograms of sub-images contributing to classification results have similar distribution. Rain drops present a high number of contiguous vertical edges. This shortcoming is sorted out using proposed method. As is shown in Figure 3(a) changing regions are marked in the binary image with white pixels and the rest of it, i.e. the static background or foreground is marked with black pixels. This information was extracted from the video sequence depicted in Figure 3(b).