

A comparison of MAE Based Image Search for Hexagonally and Regularly Structured Pixel Data

Sijing Liu^{1, a}

¹School of Electrical and Information Engineering, Chongqing University of Science and Technology, Chongqing, 401331, China

^asijingprince@foxmail.com

Keywords: Mean Average Error; Hexagonally Structured Pixel Data

Abstract. Motion tracking is an important application in video processing. By tracking motion, we can detect the presence of moving object, track and analyze the motion of object and even make prediction of the next step of a selected object. Motion tracking can be applied to facial recognition, intelligent transportation system, medical imaging and so on. Mean average error is a main step in motion tracking which set a reference block in the former frame and compare it with every block having the same size within a selected searching area in the next frame. In addition, hexagonal pixel based images have been recently investigated for image capture and processing due to their ability to represent curved structures that are common in real images better than traditional rectangular pixel based images. Therefore, we present an approach to mean average error on hexagonal images and demonstrate that accuracy is comparable rectangular pixel based images.

Introduction

In the field of image processing, square structure has been dominant for many years and mature techniques based on square structure have been developed by researchers. However, hexagonal lattices, which emerged at the same time as digital images, have attracted many scientists' interest. During the past 40 years' study, various experiments comparing performance on hexagonal structured images and square structured images have been conducted and advantages of hexagonal structured images over square structured images have been accepted by most scientists and researchers.

This paper focuses on the performance of MAE in motion tracking, which can be widely applied in video processing, in both square pixel-based images and hexagonal pixel-based images. The result will be analyzed as preliminary work for possible further study on hexagonal pixel-based images applied in video processing.

Hexagonal Structured Image

The repetition of a pattern is called tessellation. From observing the paintings and drawings from ancient civilizations, it is easily found that most of them have used regular triangular, square, and hexagonal tessellation. In fact, those three tessellations are proved to be the only three regular planar tessellations since their interior angle can exactly be divisor of 360 degrees and fill up each vertex in a plane (shown as Figure 1). Other types of polygons will result in a sensory configuration where pixel are either at different distances from their neighbors or involve gaps and overlaps.

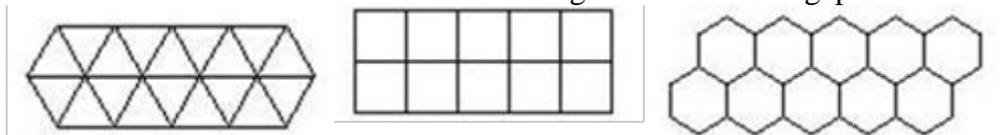


Figure 1. Three Approaches to Image Representation

Among all the three tessellations, square has held the dominant position in image representation because it is conceptually easy to develop 2D processing algorithms from the 1D case. At the same time, square lattice permits the use of natural coordinates such as Cartesian and also it supports a

separable discrete Fourier transformation (DFT). However, people got inspiration from the nature that hexagons may be the best choice to partition an image. For example, the honeycomb, known as the best 2D example of hexagonal lattice in nature, has led to the famous honeycomb conjecture which states that a hexagonal grid is the best way to divide a surface into regions of equal area with the least total perimeters [1]. The honeycomb conjecture was proposed by Pappus of Alexandria [2] and proved by mathematician Thomas C. Hales elegantly in 1999[3].

In this way, hexagonal structure has come to the attention of image processing and been researched ever since. After nearly 40 years' study, now at least six advantages of hexagonal structured images over square structured images are generally accepted by scientists [4]. For example, the angular resolution of hexagonal structure is much better than square structure which can be easily seen from representing a curved structure. Other advantages include both spatial and spectral advantages: high efficiency which was proved by Peterson and Middleton in 1962 [5], uniform connectivity, perfect isoperimetry and equidistance and high symmetry.

Mean Average Error

Motion tracking is an important application in video processing. By tracking motion, we can detect the presence of moving object, track and analyze the motion of object and even make prediction of the next step of a selected object. Motion tracking can be applied to facial recognition, intelligent transportation system, medical imaging and so on. The basic idea of motion tracking is to set a reference block in the former frame and compare it with every block having the same size within a selected searching area in the next frame. The one with the least difference would be the matching one which means the tracking object has moved to this position.

MEA in regular square pixel-based images

Suppose we have an image in which the size of the search area is $m * n$ and the size of then reference block is $i * j$. In the search area, we take one $i*j$ block as a unit. From the first unit whose start point is the start point of the search area to the last unit whose end point is the end point of the search area, calculate every unit's mean average error compared with the reference block. The method to calculate the mean average error is very straightforward. For each pixel in the reference block, subtract the value of its corresponding pixel in the assessed block within the search area and take absolute value of the difference. Sum every pair of pixels' difference and divide it by the size of the reference block. The formula to calculate the mean average error positioned at (x, y) is shown below:

$$MAE(x, y) = (1/i * j) * abs((\sum_{k=0}^{i-1} \sum_{l=0}^{j-1} A(x + k, y + l) - R(k, l))) \quad (1)$$

Here we store the $i * j$ reference block and the $m*n$ search area in two separate image which can be more convenient to do the calculation. $A(x+k,y+l)$ represents the currently assessed block of which the start point is (x, y) . $R(k,l)$ presents the reference block which always starts from its origin point $(0,0)$ and ends at $(i-1, j-1)$. Find out the smallest MAE and the corresponding position of this block. This is the current position of the tracking object.

MAE in hexagonal pixel-based images

In hexagonal pixel-based images, the basic idea of MAE algorithm is the same as the one in regular square pixel-based images. The only difference between them is the different structure of square structure and hexagonal structure. The process of MAE in hexagonal pixel-based images is the same as that in regular square pixel-based images except the method to calculate mean average error. As can be seen from figure 6, the pixels in even row and the pixels in odd row have different pattern from each other. So the mean average error in hexagonal pixel-based images consists of two sub mean average error – one based on the even row pixels and the other based on the odd row pixels. Thus we have three formulas to calculate the mean average error in hexagonal pixel-based images.

$$MAE_{even-row}(x, y) = \left(\frac{1}{\binom{i}{8} \binom{j}{14}}\right) * abs((\sum_{k=0}^{i/8} \sum_{l=0}^{j/14} A_{even-row}(x+k, y+l) - R_{even-row}(k, l))) \quad (2)$$

$$MAE_{odd-row}(x, y) = \left(\frac{1}{\binom{i-4}{8} \binom{j-7}{14}}\right) * abs((\sum_{k=0}^{(i-4)/8} \sum_{l=0}^{(j-7)/14} A_{odd-row}(x+k, y+l) - R_{odd-row}(k, l))) \quad (3)$$

$$MAE(x, y) = MAE_{even-row}(x, y) + MAE_{odd-row}(x, y) \quad (4)$$

Here we store the $i*j$ reference block and the $m*n$ search area in two separate images which can be more convenient to do the calculation. Based on the hexagonal pixel structure, the number of even columns in the reference block is $i/8$, the number of odd columns is $(i-4)/8$, the number of even rows is $j/14$ and the number of odd rows is $(j-7)/14$. $A_{even-row}(x+k, y+l)$ and $A_{odd-row}(x+k, y+l)$ represent the currently assessed block of which the start point is (x, y) . $R_{even-row}(k, l)$ and $R_{odd-row}(k, l)$ presents the reference block which always starts from its origin point $(0,0)$ and ends at $(i-1, j-1)$.

Results

Currently, there are no hexagonal camera sensors commercially available and therefore hexagonal pixel-based images are created via resampling. The resampling technique we use is the 56 sub-pixel approach in [6] which is based on the technique of Wuthrich [7]. In order to create a sub-pixel effect to enable the sub-pixel clustering, each pixel of the original rectangle pixel based image is represented by a $7*7$ pixel block of equal intensity in the new image. This creates a resized image of the same resolution as the original image with the ability to display each pixel as a group of $n*n$ sub pixels and limits the loss of image resolution. Another motivation for image resizing is to enable the display of sub pixels, which is not otherwise possible. With this structure, a cluster of 56 sub pixels in the new image, closely representing the shape of a hexagon, can be created that represents a single hexagonal pixel in the resized image. At the same time, in order to get a good visualization, we resize the original image by 7 times which then can be as the same size as the resampled hexagonal pixel-based image.

For comparison, the visual responses from each of the two kinds of images are illustrated in Figure 2, Figure 3.

The first two figures (Figure 2 and Figure 3) show the MAE algorithm applied in a square pixel-based image and a hexagonal pixel-based image which is resampled from the former. Figure 2 displays the situation that the object in the reference block also exists in the search area. It can be seen both the two kinds of images perform well and the exact reference block is found in the search area. While in Figure 3, the search area does not contain the searched object but has a similar object. In this situation, both of these two kinds of images successfully identified the similar object. But there is a tiny difference between them. Although MAE in both of these two kinds of images can identify the searched object in the search area, the exact positions of the searched reference block in the search area are not always the same. As can be seen from Figure 3(b) and 3(c), the searched reference block in the hexagonal pixel-based image is a little higher than that in the square pixel-based image.

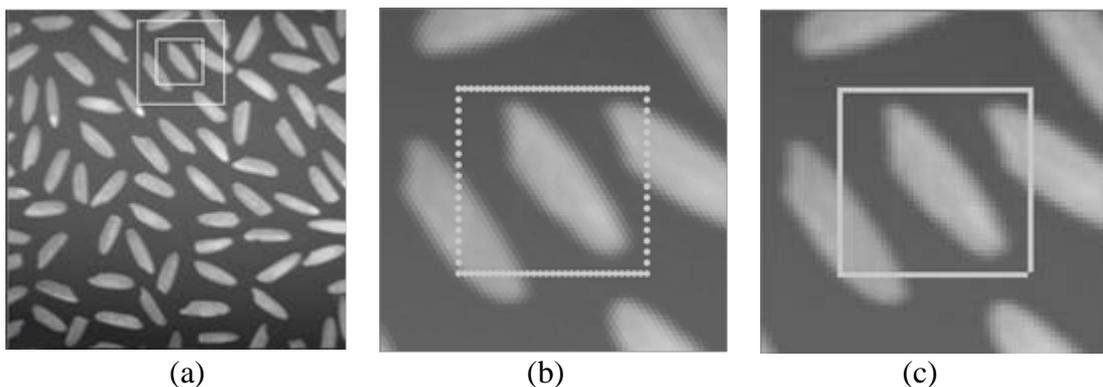


Figure 2. Searching for an object which exists in the search area

(a)The test image in which red rectangle represents reference block and white rectangle represents search

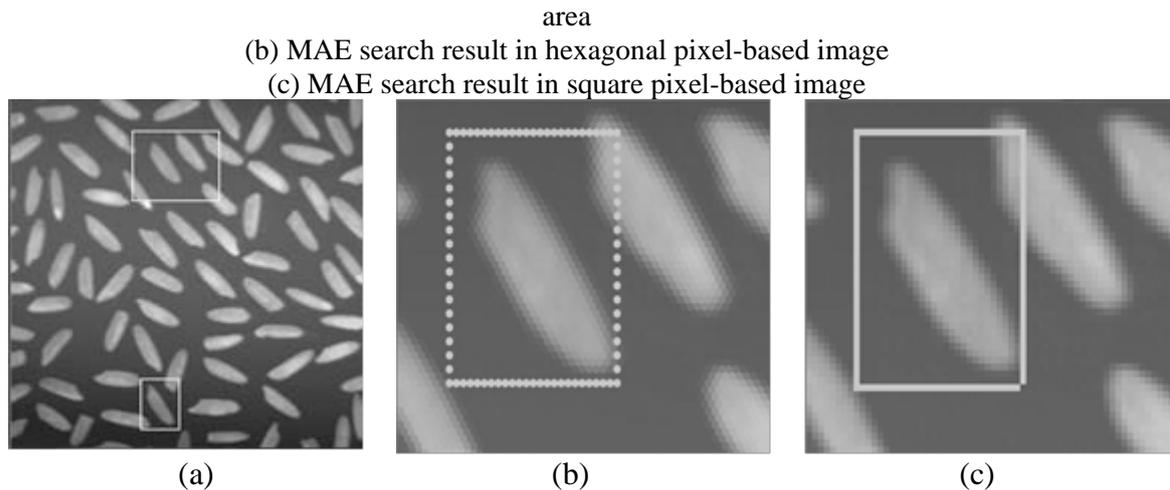


Figure 3. Searching for an object in a search area which contains a similar object
(a) The test image in which red rectangle represents reference block and white rectangle represents search area
(b) MAE search result in hexagonal pixel-based image
(c) MAE search result in square pixel-based image

Summary

Motion tracking plays an important role in video processing which can be widely applied to various fields in human society. Since the current camera sensors generate square pixel-based images, video processing techniques based on square pixel-based image have been developed maturely. However, with people's rising interests toward hexagonal pixel-based image and its application in image processing and real-time vision, kinds of experiments based on hexagonal pixel-based images have been conducted to compare the performance with traditional square pixel-based images. In this project, we present an approach to implement MAE algorithm in hexagonal pixel-based images. Based on the results of the project, it can be concluded that the accuracy of detected effective pixels in hexagonal pixel-based image is high and the performance of hexagonal pixel-based images is more stable than that of square pixel-based image which suggests that hexagonal pixel-based images may have a future application in video processing. However the limitation of this project is that we used fixed image instead of moving images. Since there are no commercial camera sensors currently generate hexagonal pixel-based images, the time spent on resampling will be the biggest obstacle in video processing based on hexagonal pixel-based images.

Reference

- [1] Wikipedia, http://en.wikipedia.org/wiki/Honeycomb_conjecture, retrieved on 17th May 2011.
- [2] Weisstein, Eric W. "Honeycomb Conjecture", Mathworld, <http://mathworld.wolfram.com/HoneycombConjecture.html> , Retrieved 17th May 2011.
- [3] Hales, Thomas C. "The Honeycomb Conjecture". *Discrete and Computational Geometry* 25: 1-22(2001).
- [4] Xiangjian He, Wenjing Jia, *Hexagonal structure for Intelligent Vision*, p. 53-55.
- [5] Peterson, D. P. and D. Middleton, *Sampling and Reconstruction of Wave-Number-Limited Functions in NDimensional Euclidean Spaces*. *Information and Control*, 1962.5: p. 279-323.
- [6] Middleton, L. and Sivaswamy, J., "Hexagonal Image Processing; A Practical Approach", Springer 2005.
- [7] Wuthrich, C.A, Stucki, P. "An algorithmic Comparison between Square- and Hexagonal-based Grid", *CVGIP: Graphical Models and Image Processing*, Vol. 53, pp. 324-339, 1999.