УДК 004.02 DOI: 10.31673/2412-4338.2020.044451

Dibrivniy O. A. *State University of Telecommunications, Kyiv*

USAGE OF CHARACTERISTIC POINTS TO REDUCE THE TIME OF PERFORMANCE OF SAD METRIC CALCULATION DURING VIDEO FLOW COMPRESSION BY MOTION COMPENSATION METHOD

Abstract: *Today, it is difficult to imagine a video system that does not use compression of the input video stream for transmission or storage, because the uncompressed size of video files can reach several hundred gigabytes. There are many different video stream compression techniques used in various codecs, however, one of the most commonly used is the motion compensation technique, which allows you to transmit video stream frames in the form of compensated inter frame difference, by finding the motion vectors of individual blocks image. Despite the fact that the above-mentioned technique has been used for the last twenty years, it still has two problem-priority places - assessments of the similarity of image blocks and the algorithm for searching blocks. This article focuses on the problem of estimating the similarity of the image, in order to reduce the time of estimating the similarity of the blocks, which in modern algorithms takes from 40 to 80% of the time of the entire process of encoding the video stream.*

The article considers the usage of metrics for estimating the similarity of images and their importance for the process of motion compensation during video stream compression, such as SSD (standard deviation) and SAD (sum of absolute differences). The temporal contribution of metrics to the video stream encoding process, in particular to the motion compensation process, is estimated. An algorithm combining classical SAD and parameter estimation based on characteristic points is proposed, which will reduce the metric calculation time for estimating block similarity. The reduction of SAD calculation time due to reduction of comparison points is estimated for three proposed comparison templates - HSAD (half SAD), TSAD (third SAD), DSAD (diagonal SAD). For the selected templates, the results of processing the test video sequence were compared with the results obtained during the processing of the video sequence by the classic SAD. The main attention was paid to the assessment of the following parameters: the relative difference SAD, the increase in the number of candidate blocks, the overlap of candidate blocks.

Keywords: *Sum of absolute differences, motion compensation, inter-frame difference, characteristic points, video stream compression.*

Дібрівний О.А. *Державний університет телекомунікацій, Київ*

ВИКОРИСТАННЯ ХАРАКТЕРНИХ ТОЧОК ДЛЯ ЗМЕНШЕННЯ ЧАСУ ВИКОНАННЯ ОБРАХУНКУ МЕТРИКИ SAD ПІД ЧАС СТИСНЕННЯ ВІДЕОПОТОКУ МЕТОДОМ КОМПЕНСАЦІЇ РУХУ

Анотація: *На сьогоднішній важко собі уявити відео-систему яка не використовує стиснення вхідного відео-потоку для передачі або зберігання, оскільки нестиснений розмір відео-файлів може досягати кількох сотень гігабайт. Існує багато різних методик стиснення відео-потоку, які використовуються в різного роду кодеках, тим не менш одною з найбільш часто використовуваних є методика компенсації руху, яка дозволяє передавати кадри відео-потоку у вигляді скомпенсованої міжкадрової-різниці, за рахунок знаходження векторів руху окремих блоків зображення. Не дивлячись на те, що вищезгадану методику використовують протягом останніх двадцяти років, в ній залишається два проблемно-пріоритетних місця – оцінки схожості блоків зображення та алгоритм перебору блоків. Дана стаття зосереджена на розгляді проблеми оцінки схожості зображення, з метою скорочення часу оцінки схожості блоків, яка в сучасних алгоритмах займає від 40 до 80% часу всього процесу кодування відео-потоку.*

В статті розглянуто використання метрик оцінки схожості зображень та їх важливості *для процесу компенсації руху під час стиснення відео-потоку, такі як SSD(середнє квадратичне відхилення) та SAD(сума абсолютних різниць). Оцінено часовий вклад метрик в процес кодування відео-потоку, в особливості в процес компенсації руху. Запропоновано алгоритм що поєднує класичний SAD та оцінку параметрів на основі характерних точок, що дозволить знизити час обрахунку метрики для оцінки схожості блоків. Оцінено зменшення часу виконання обрахунку SAD за рахунок зменшення точок порівняння, для трьох запропонованих шаблонів порівняння – HSAD(half SAD), TSAD(third SAD), DSAD(diagonal SAD). Для обраних шаблонів проведено порівняння результатів обробки тестової відео-послідовності з результатами отриманими при обробці відео-послідовності класичним SAD. Основну увагу було приділено оцінці наступних параметрів: відносній різниці SAD, збільшенню кількості блоків кандидатів, перекриття блоків кандидатів.*

Ключові слова: *Сума абсолютних різниць, компенсація руху, міжкадрова різниця, характерні точки, стиснення відео-потоку.*

Дибривный О.А. *Государственный университет телекоммуникаций, Киев*

ИСПОЛЬЗОВАНИЕ ХАРАКТЕРНЫХ ТОЧЕК ДЛЯ УМЕНЬШЕНИЯ ВРЕМЕНИ ВЫПОЛНЕНИЯ РАСЧЕТА МЕТРИК SAD ПРИ СЖАТИИ ВИДЕОПОТОКА МЕТОДОМ КОМПЕНСАЦИИ ДВИЖЕНИЯ

Аннотация: *На сегодняшний трудно себе представить видеосистему, которая не использует сжатие входящего видеопотока для передачи или хранения, поскольку несжатый размер видеофайлов может достигать нескольких сотен гигабайт. Существует много различных методик сжатия видеопотока, которые используются в разных кодеках, тем не менее одной из наиболее часто используемых является методика компенсации движения, которая позволяет передавать кадры видеопотока в виде скомпенсированной межкадровой разницы, за счет нахождения векторов движения отдельных блоков изображение. Несмотря на то, что вышеупомянутую методику используют в течение последних двадцати лет, в ней остается два проблемно-приоритетных места – оценки сходства блоков изображения и алгоритм перебора блоков. Данная статья сосредоточена на рассмотрении проблемы оценки сходства изображения с целью сокращения времени оценки сходства блоков, которая в современных алгоритмах занимает от 40 до 80% времени всего процесса кодирования видеопотока.*

В статье рассмотрено использование метрик оценки сходства изображений и их важности для процесса компенсации движения при сжатии видеопотока, такие как SSD (среднее квадратическое отклонение) и SAD (сумма абсолютных разниц). Оценен временной вклад метрик в процесс кодирования видеопотока, в особенности в процесс компенсации движения. Предложен алгоритм сочетающий классический SAD и оценку параметров на основе характерных точек, что позволит снизить время расчета метрики для оценки сходства блоков. Оценено уменьшение времени выполнения расчета SAD за счет уменьшения точек сравнения, для трех предложенных шаблонов сравнения - HSAD (half SAD), TSAD (third SAD), DSAD (diagonal SAD). Для избранных шаблонов проведено сравнение результатов обработки тестовой видеопоследовательности с результатами полученными при обработке видео последовательности классическим SAD. Основное внимание было уделено оценке следующих параметров: относительной разницы SAD, увеличению количества блоков кандидатов, перекрытия блоков кандидатов.

Ключевые слова: *Сумма абсолютных разниц, компенсация движения, межкадровая разница, характерные точки, сжатие видеопотока.*

Introduction

Motion compensation in a video stream is a process of obtaining information about the nature and parameters of the so-called "optical" two-dimensional movement of objects (in the plane of the frame) by the available video information. The use of this information can significantly increase the degree of compression of video compression algorithms and provides additional capabilities to video processing algorithms. The motion compensation module is an integral part of almost all modern video codecs, video surveillance systems and the highest quality noise reduction algorithms. Currently, the most popular is the block approach to motion compensation, in which each frame is divided into equal size blocks (usually 16x16 pixels, although modern systems use blocks of 32x32 and 64x64 pixels) and for each block is the search for the closest to it (in a certain metric space) block of the same size in the previous frame [1]. Since the calculation of this metric for two arbitrary blocks usually requires a number of operations of the order of the square of the block size, the main task is to develop a strategy for finding the minimum of a given metric, which would require a minimum number of calculations.

1. Description of initial data.

When calculating the results of the proposed algorithm, we will use 2 sets of data, with a frame size of 1920x1080 pixels. The first data set will be two matrices of according size, with randomly generated cell values in the range from 0 to 255, which corresponds to the dimension of the color component for 8-bit format. The second set will be a video stream with a length of 1 second, frequency of 29.97 frames per second, in MPEG4 format (H264) and color format YUV 4: 2: 0, color depth - 8 bits. To get the best result of the illustrative evaluation we use a full search of the area ± 32x± 32 pixels around the processed block, which will lead to 4096 calculations for each of the selected metrics for the block (more than 8 thousand blocks per frame) and as a result of metric calculations for 2 frames. The purpose of this article is to form a method for estimating the similarity of images, which will optimize the speed of algorithmic calculations with minimal loss in quality/truth of the information of the processed images.

2. Algorithm for the sum of absolute differences.

SAD is an extremely fast indicator due to its simplicity. In essence, this is the simplest of the possible metrics and is calculated by taking the absolute difference between each pixel in the input macroblock and the corresponding pixel in the block used for comparison. These differences are summed to create a simple block similarity metric. One of the main advantages of the SAD metric in addition to the speed of the addition operation is the easy process of parallelization, because each pixel is processed separately, which allows you to implement an algorithm with system instructions such as ARM NEON or x86 SSE2. Nevertheless, the calculation of SAD for video sequence frames can take from 40 to 80% of the total coding time of the video stream. The accuracy of the SAD algorithm is usually insufficient to form the correct motion vectors, and in most cases, after finding the candidate blocks, the final choice of the block is made using other, slower, but more accurate indicators that better take into account human perception, such as absolute converted differences (SATD), sum of squares differences (SSD) and Rate-distortion optimization (RDO) [2].

To calculate the SAD for two frames A and B, we can write the following formula:

$$SAD(A,B) = \sum_{i=1}^{n} |a_i - b_i| \tag{1}$$

where $A = \{a_1,...a_n\}$ and $B = \{b_1,...b_n\}$ - two sets of pixel data corresponding to frames. When compressing a video stream, the sum of the absolute differences is most often calculated for macroblocks of 16 by 16 pixels. Then from the obtained SAD, within some deviation, a set of candidate vectors is formed, which is further processed by one of the slower methods.

One way to reduce the number of operations, and as a consequence of the execution time of the algorithm is to use templates for comparison or so-called comparison of characteristic points. During this process, instead of taking the difference between each pixel of the current macroblock and the corresponding pixel from the macroblock of the reference frame, a template is selected according to which part of the pixels will be ignored. Given that the image is essentially a matrix, the size of which coincides with the expansion of the image and the input macroblock is a matrix of 16x16, you can link the template to the visual image of the matrix. The templates used in the calculation of the SAD algorithm are presented in Fig.1. where the pixels of the macroblock to be searched are marked in black.

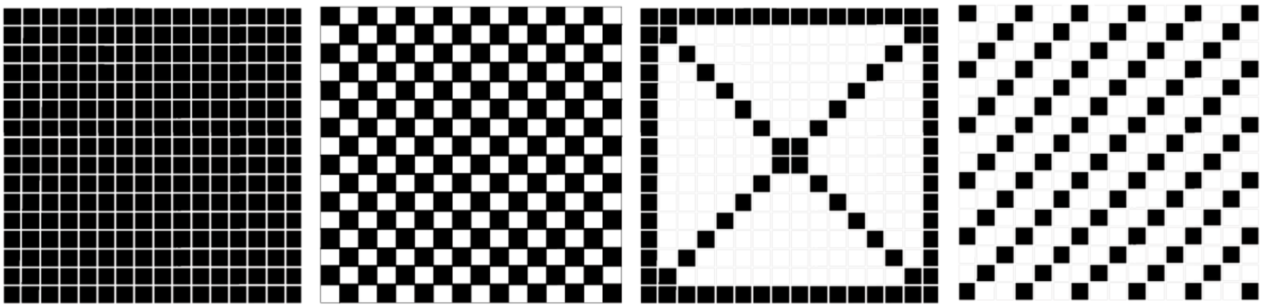

Fig.1. SAD execution templates, a) full search, b) search on the main diagonals and edge points, c) "chess" search (search of every second pixel) c) search of every third pixel

For convenience of the description of the chosen templates we will mark them as follows. full search - full SAD (FSAD), "chess" search - half SAD (HSAD), diagonal - diagonal SAD (DSAD), search every third pixel - third SAD (TSAD).

The reduction in the number of operations and as a consequence of the reduction of the calculation time for the selected templates are presented in table 1.

Table 1

Decrease the number of operations for the selected search patterns.

	HSAD	DSAD	TSAD
Operations number	128	88	86
Decrease in operations number relative to FSAD	128	168	170
%reduction in operations number relative to FSAD	50	65.6	66.4

As can be seen from the table due to the usage of comparison templates, we obtained a significant reduction in the number of operations and as a consequence reduction of the algorithm execution time, which even when operations are paralleled can reach a similar gain in time. This raises the question of how much, reduction in the number of result points will affect the accuracy of the selection of candidate blocks. To begin with, we will evaluate randomly generated data (it should be noted that randomly generated data is not subject to effective compression, but for us the main goal is to show the effectiveness of finding similar blocks). To do this, run an algorithm for selecting candidate vectors on a matrix of 1920x1080 with a search area of 64 pixels, for all four search templates, then compare the results obtained using FSAD with each of the 3 templates, and determine the relative difference of absolute differences. The results of calculations for two matrices with a search area of ± 32x ± 32 pixels can be seen in Fig.2.

From the table 2 we can conclude that the DSAD template with more control points than TSAD shows a worse result due to the relatively low distribution density of control points in the middle of the matrix and the relative simplicity of the uniqueness of the pattern. TSAD with a smaller number shows better results because the diagonal arrangement allows you to more accurately adjust to random matches of pixel values.

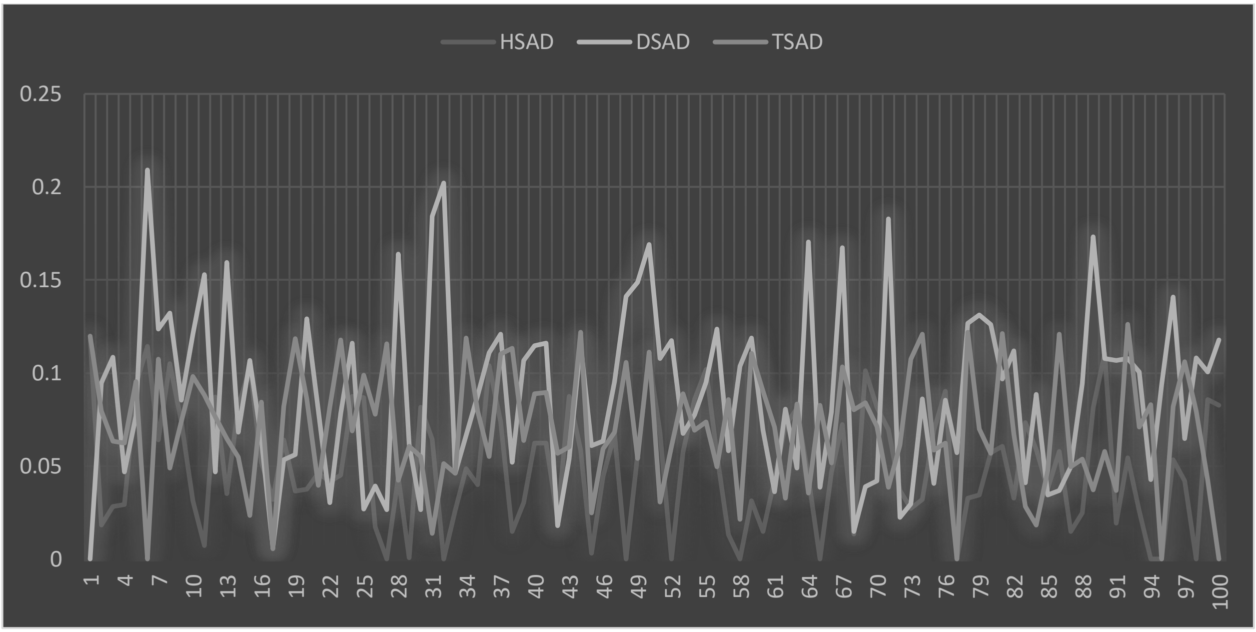

Fig.2. Distribution of values of relative difference between FSAD and the offered comparison templates

Table 2

Values of average and maximum relative deviations for the selected comparison templates, for the calculation of random data

	HSAD	DSAD	TSAD
Maximum relative deviation	0,111	0,209	0,129
Mean relative deviation	0,053	0,098	0,076

Now consider a real set of video stream data with the characteristics described above. Fig. 3 shows the first two frames of the processed video stream.

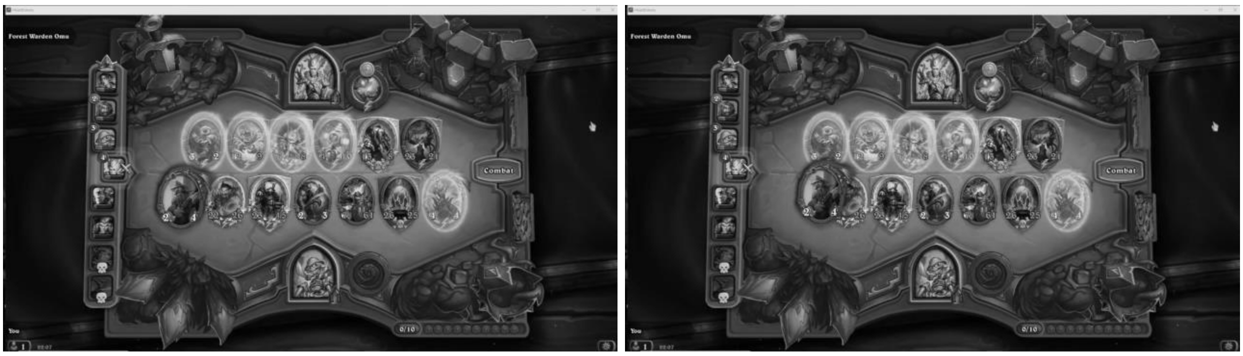

Fig.3. The first 2 frames of the processed video stream

As a result, we obtain a completely different picture of the values of the relative deviation of the sum of the absolute differences (Table 3).

In what follows, we will immediately reject the use of the DSAD template due to its relatively high deviation in the sum of absolute differences, with the same reduction in the number of operations compared to TSAD, which is associated with large void areas (areas not taken to calculate pixels).

Table 3

Values of average and maximum relative deviations for the selected comparison templates, for the calculation of real data

	HSAD	DSAD	TSAD
Maximum relative deviation	0,048	0,112	0,052
Mean relative deviation	0,025	0,063	0,028

When calculating real data, another important indicator of the effectiveness of the chosen algorithm for estimating the similarity of blocks is the number of candidate blocks for each block. Moreover, this number is usually the largest in the homogeneous areas of the frame, which remain unchanged when changing the frame. For the selected video sequence, the number of candidate blocks for immutable areas when using the SAD algorithm can reach several tens (for the test video sequence, the maximum number of such blocks reached 14 - due to the low entropy of the image). However, to calculate such areas in practice, you can take any of the blocks of candidates, as their SAD will be equal to 0, or quite close to it. That is, we are interested in those blocks for which SAD will be greater than some value, if neglected, we will not have a visual change in the image that can be caught by the human eye [3].

As for the blocks that interest us the most, these are the blocks in which there is an inter frame difference. The absolute uncompensated difference of the first two frames of the video sequence (Fig.4).

Fig.4. Absolute uncompensated difference of the first two frames of the video sequence

In the part of the image with the inter frame difference, we will select an area of 10x10 blocks (160x160 pixels) in the second frame of our video sequence (the beginning of the area at point T (576, 512), the results of FSAD, HSAD and TSAD calculation which we will evaluate to obtain information on the number of candidate vectors (Fig. 5).

So the distribution of the number of candidate blocks of size 16x16, for the selected SAD algorithms are presented in Fig. 6.

As we can see from Fig.6 the use of HSAD and TSAD has led to an increase in the number of candidate blocks for the wanted blocks, due to which the refinement time of candidate blocks may increase. Also important to us are the points where SAD gives more candidate vectors than TSAD does.

Fig.5. Test area for evaluating the work of the proposed algorithms

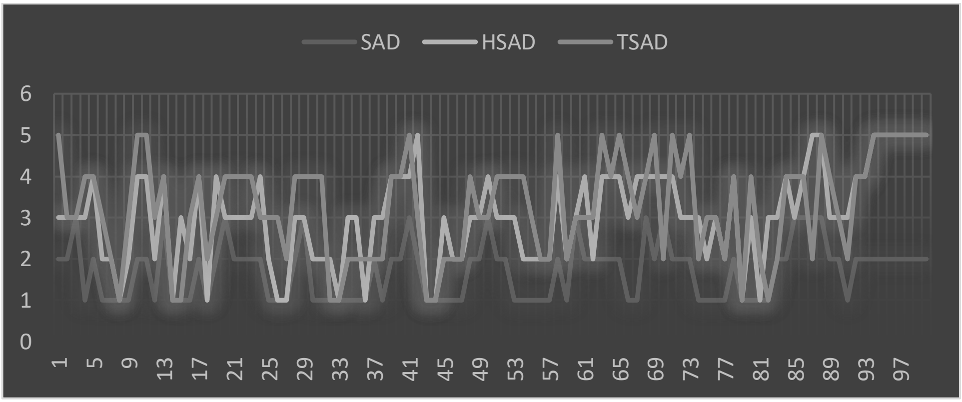

Fig. 6. The number of candidate vectors for 100 blocks in the middle of the calculation area

This result can have two consequences:

1. Reducing the number of blocks of candidates due to the uniqueness of the pattern (selection of the best possible options due to the exact match of pixels).

2. Reducing the number of candidate blocks by skipping blocks in which the SAD value fits under the defined and is best for a particular block.

As a consequence of reducing the number of candidate blocks, it is advisable to count the overlap of the received candidate blocks obtained using TSAD and HSAD, relative to those received by SAD and the average number of candidate blocks for each template for the selected area and the whole image (Table 4):

Table 4

General characteristics of the number and overlap of candidate blocks

	SAD	HSAD	TSAD
The average number of candidate blocks for the calculation area	1,83	3,21	3,43
The average number of candidate blocks for the entire image	14.31	16,99	17,67
Overlap of blocks obtained by SAD and the selected algorithm for the calculation area	_____	97,92%	97,14%

As can be seen from the table, for the calculation area the number of candidate vectors increased by 43% for HSAD and 47% for TSAD for the selected area, or 15.8% and 19.9% for the entire search area (the decrease in the average number of blocks across the area is due to in some cases, the proposed algorithms show the best selective search due to the uniqueness of the pattern, for areas where the number of blocks of candidates is growing), which can lead to a corresponding increase in time at the stage of refining blocks, which takes 5-15% of video stream encoding time. However, the gain in time still remains quite noticeable, estimates of the execution time of the full algorithm for determining the blocks of candidates will be given below.

Another consequence of the use of HSAD and TSAD algorithms is the loss of 2.08% and 2.86% of candidate blocks obtained through the use of SAD, however, since SAD itself does not give a perfect match of the blocks, we can ignore the error of this level. It is also worth noting that the greatest loss of candidate blocks will take place in areas with minimum entropy, while raise of block candidates count will take place in areas with maximum entropy - on the border of objects and a sharp transition of color [4]. This can cause artifacts to appear on the boundaries of objects such as blurring or border distortion. Another consequence of using selected comparison templates may be a change in image brightness.

Conclusion

The algorithms proposed in the article allow to reduce the time for calculating the metric similarity of SAD images, through the use of comparison templates through characteristic points. Reducing the number of points for calculating SAD by 50% leads to an increase in the number of vectors of candidates in the area with an inter frame difference of 43% and overlap of 97.92%, reducing the number of calculation points by 66.4% - increasing the number of candidate blocks is 47% with overlapping 97.14 %. From the above results we can conclude. what is more promising is usage of the TSAD template. The main disadvantages of using this template are the loss of some data and increase in block candidates count, which can lead to the appearance of artifacts and complicate the process of refining the required blocks among the blocks of candidates. It is worth noting that the loss of candidate blocks will be more acute in homogeneous areas, while the increase in the number of blocks at the boundaries of objects, so it is worth considering the algorithms for selecting edges as a supplement to the metrics of refining blocks.

References

1. Chenetal T.C. (2006) Analysis and Architecture Design of an HD720p 30 Frames/s H.264/AVC Encoder. *IEEE Trans. Cir. Syst. Video Tech.*, vol. 16, no. 6. 673-688pp.

2. Deng L., Gao W. (2005) An efficient hardware implementation for motion estimation of AVC standard. *IEEE Trans. Consumer Electron.*, vol. 51, no. 4, 1360-1366pp.

3. Duanmu X. Q., Duanmu C. J., Zou C. R. (2005) A multilevel successive elimination algorithm for block matching motion estimation. *IEEE Trans. Image Process*, vol. 9, no. 3, 501-504pp.

4. Liu J., Yuan L., Xie X., (2016) Hardware-Oriented Adaptive Multiresolution Motion Estimation Algorithm and Its VLSI Architecture. *IEEE Trans. Image Process,* vol.78, no.1, 4799-5341pp.