E-JUST

Fast Fractional-Pixel Motion Estimation using Lagrangian-Based Error Surface Interpolation

Emad Badry, Ahmed Shalaby, Mohammed S. Sayed

By/Emad Badry
(emad.mahmoud@ejust.edu.eg)

The fifth IEEE Global Conference on Signal and Information Processing (GlobalSIP2017)

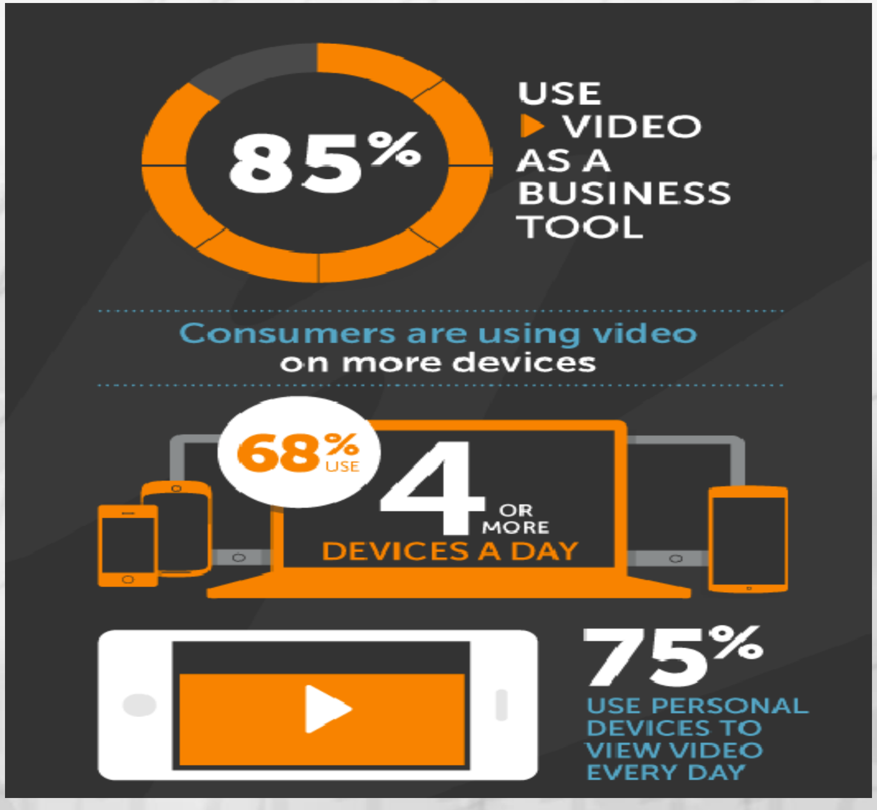
Presentation Outlines

- **Part I:** Overview of current video coding technology .
- **Part II:** High Efficiency Video Coding (HEVC)
- **Part III:** Interpolation-Free Pixel Motion Estimation.
- **Part IV** Proposed Algorithm.
  - Mathematical Model.
  - Algorithm Steps.
  - Computational Complexity Analysis.
  - Results

Part I: Overview of current video coding technology .

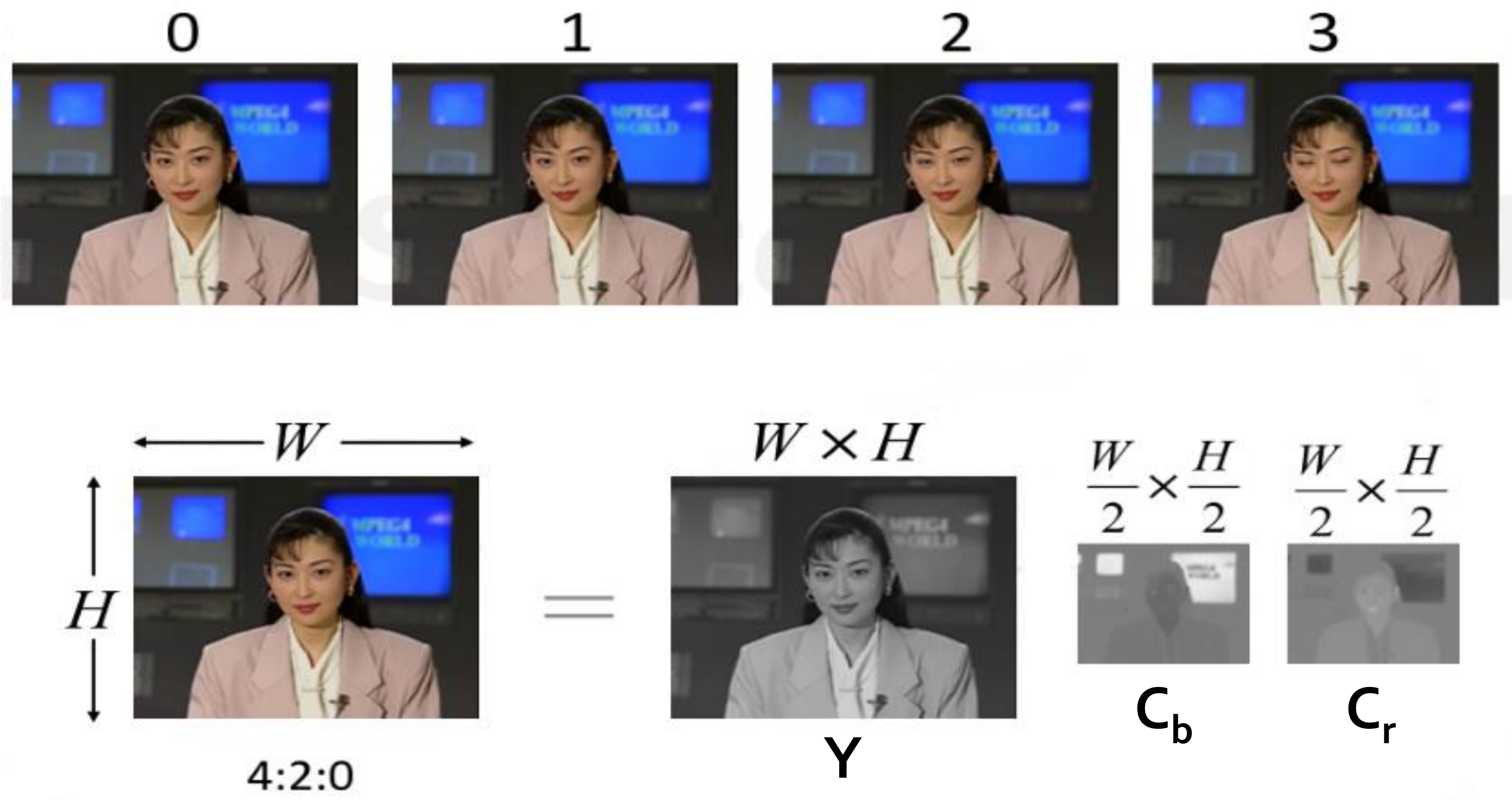
Overview

- **Growing Demand for Video**



Overview

- **Digital Video**



Overview

Video Compression Basics

- Compression is achieved by removing redundant information from the video sequence
- Types of redundancies in video sequences
  - Spatial redundancy.
  - Perceptual redundancy.
  - Spectral redundancy.
  - Temporal redundancy.

Part II: High Efficiency Video Coding (HEVC).

HEVC

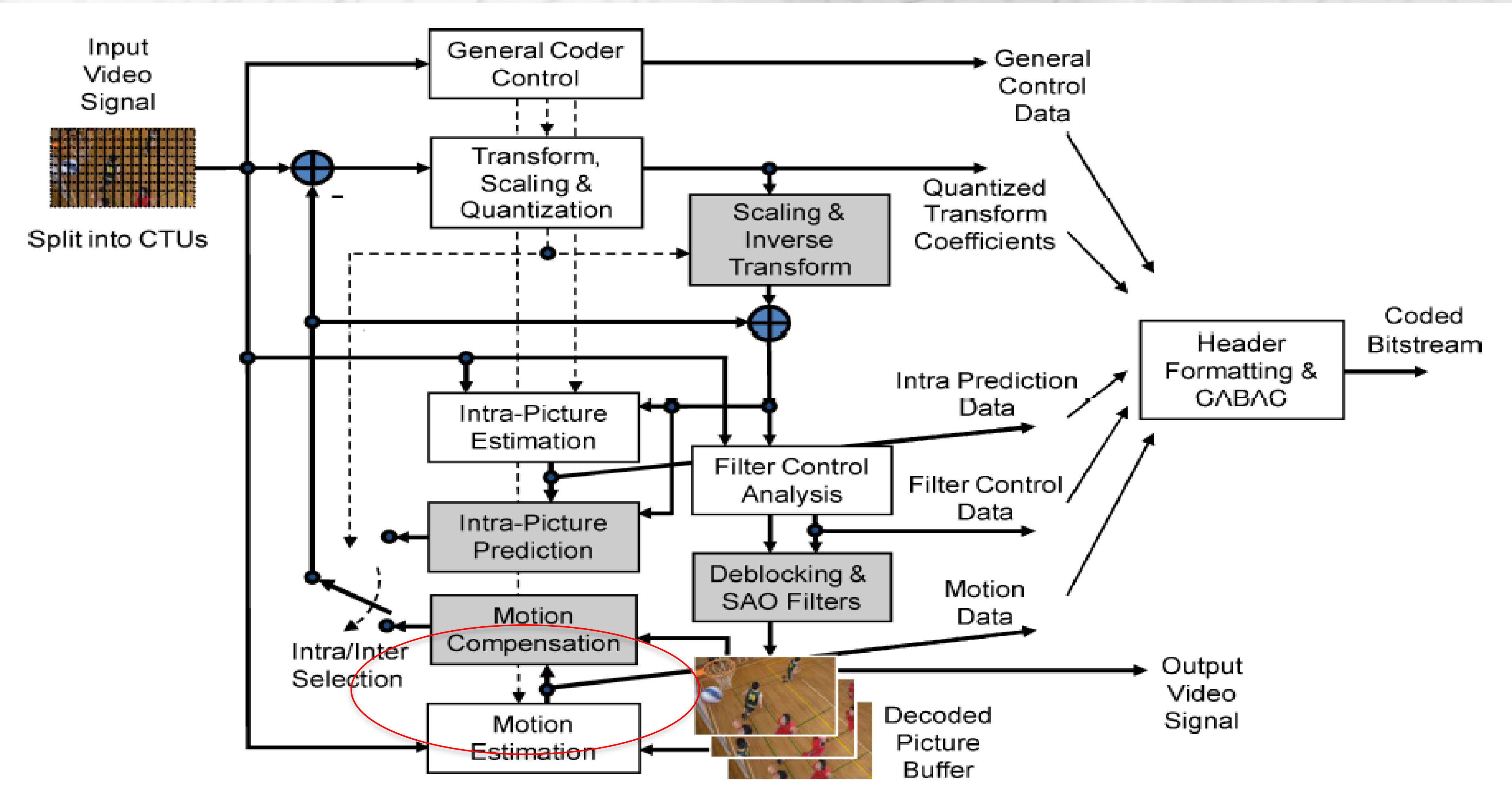
- The High Efficiency Video Coding (HEVC) is the new video coding standard that was jointly developed by the two standardization organizations, ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) .

- HEVC was developed to increase the compression efficiency by reducing the bit rate 50% with respect to the H.264/AVC standard.

- HEVC has been designed to address essentially all existing applications of H.264/MPEG-4 AVC and to particularly focus on two key issues: increased video resolution and increased use of parallel processing architectures [1].

HEVC

- Typical HEVC video encoder

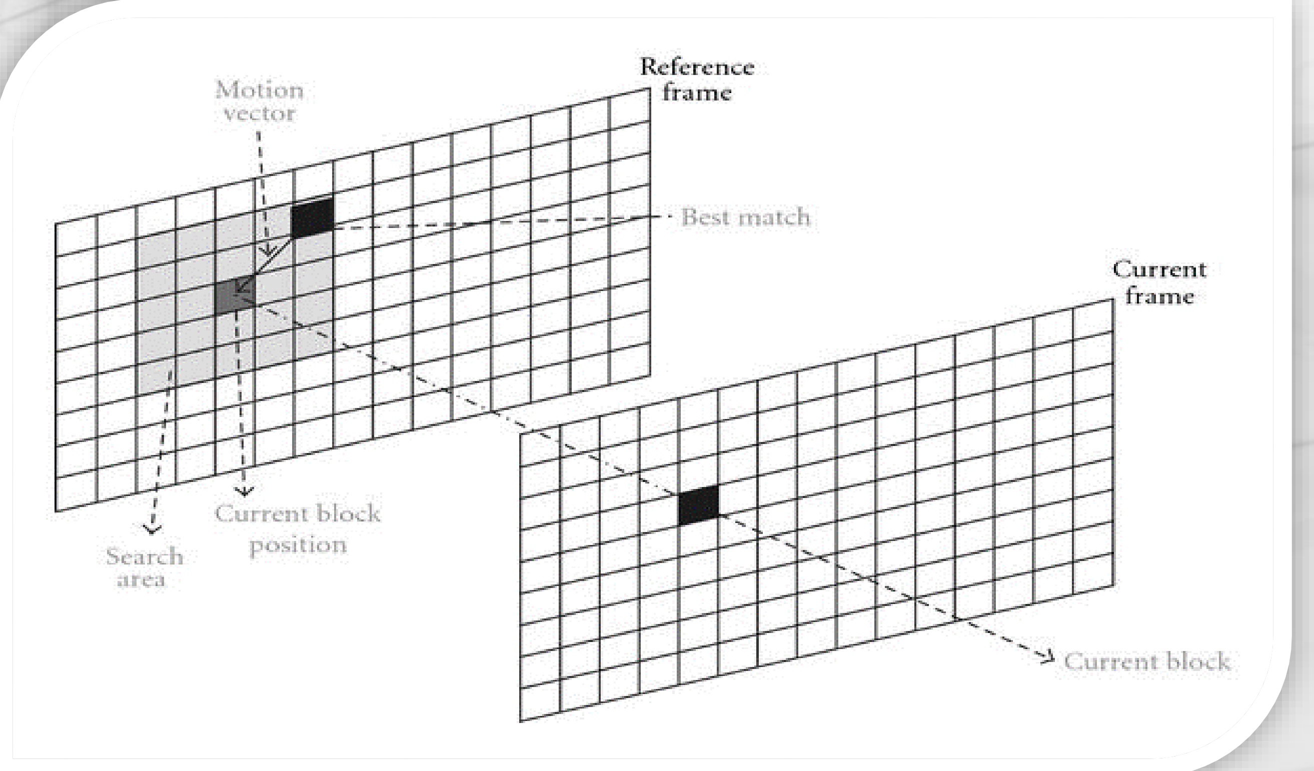


[1].

HEVC

Motion estimation is divided into two steps:

- The first step is to estimate the best integer pixel location to provide the Integer pixel Motion Vector (IMV). This step is performed by Integer pixel Motion Estimation (IME) unit applying search strategies such as full search or fast search.



Motion estimation process [3]

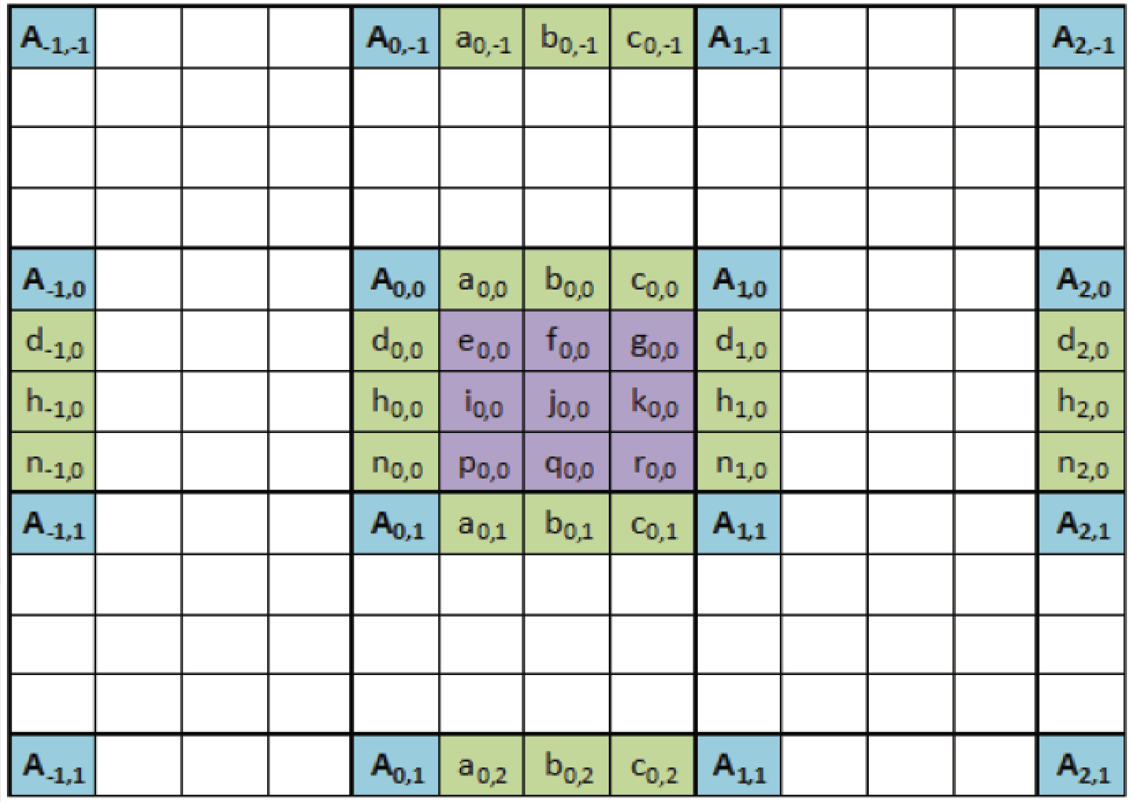
HEVC

- The second step is to perform Fractional pixel Motion Estimation (FME) around the estimated best integer pixel location in order to find the sub-pixel (i.e. fractional pixel) location with the minimum matching error for more performance improvement.

A- Interpolation Process

8-tap filters and 7-tap filters

Type	Coefficients							
A type	-1	4	-10	58	17	-5	3	0
B type	-1	4	-11	40	40	-11	4	-1
C type	0	1	-5	17	58	-10	4	-1



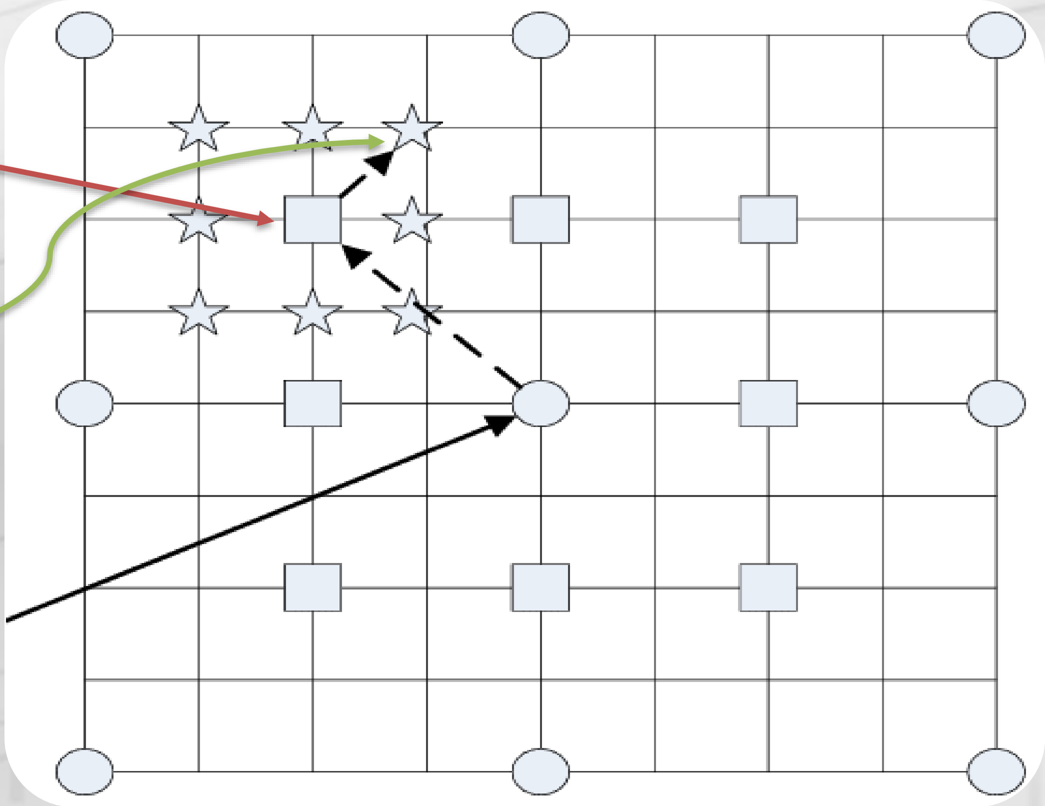
E-JUST

B- Best fractional pixel estimation

After the interpolation process, the best fractional location can be estimated as follow:

- First, find the best half-pixel location by examining the eight half-pixel positions around the best integer position found by the IME unit.

- Second, the eight quarter-pixel positions are examined to estimate the best sub-pixel position with quarter-pixel accuracy



This method suffers from high computational complexity, memory requirements and large encoding time

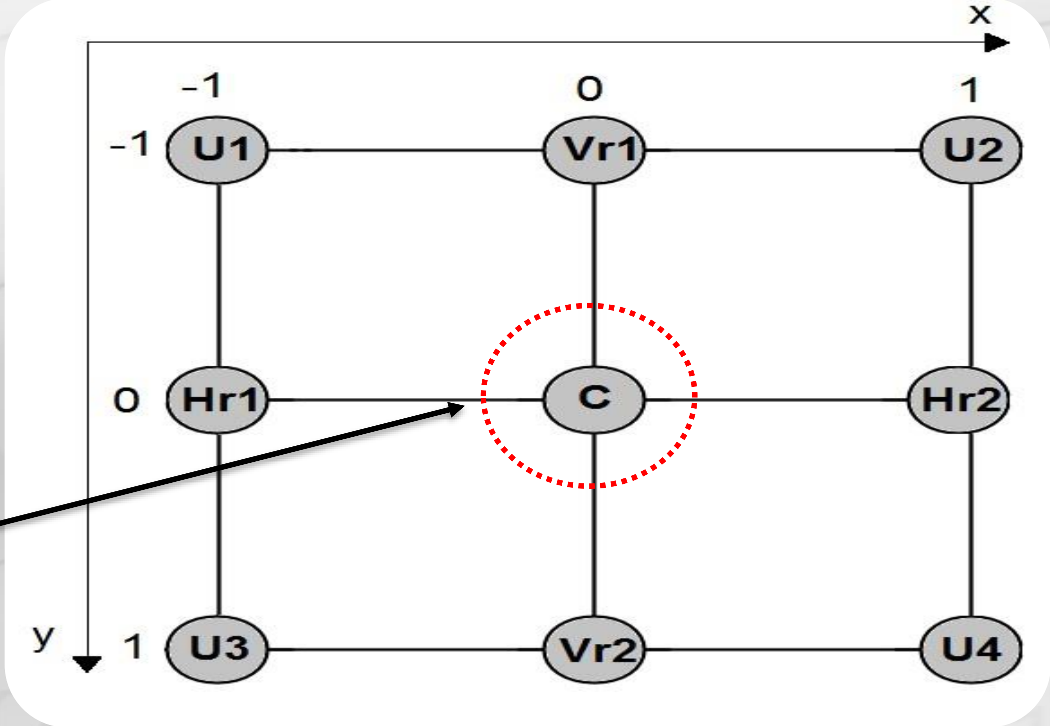
Part III: Interpolation-Free Pixel Motion Estimation.

Interpolation-Free Pixel Motion Estimation

E-JUST

- Different approaches were presented to reduce the FME encoding complexity by estimating the matching error at fractional-pixel positions directly without interpolation or matching error calculation processes.

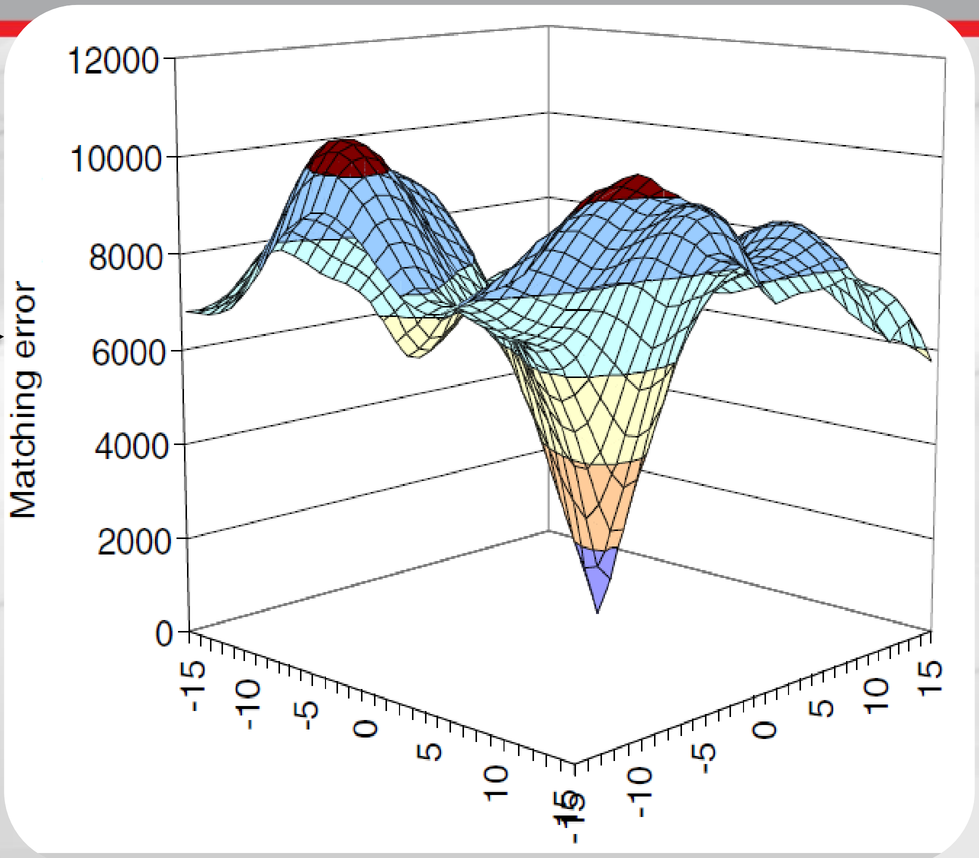
- The best integer-pixel locations can be modeled mathematically using the matching error values at eight neighboring integer pixel locations surrounding the best one.



Best integer location

Interpolation-Free Pixel Motion Estimation

Several mathematical models for the error surface have been proposed using 2-D paraboloid functions



- Several mathematical models for the error surface have been proposed using 2-D paraboloid functions, including: 9-terms model [2], 6-terms model [3] and 5-terms model [4].

$$f_9(x,y) = Ax^2y^2 + Bx^2y + Cxy^2 + Dx^2 + Exy + Fy^2 + Gx + Hy + I$$

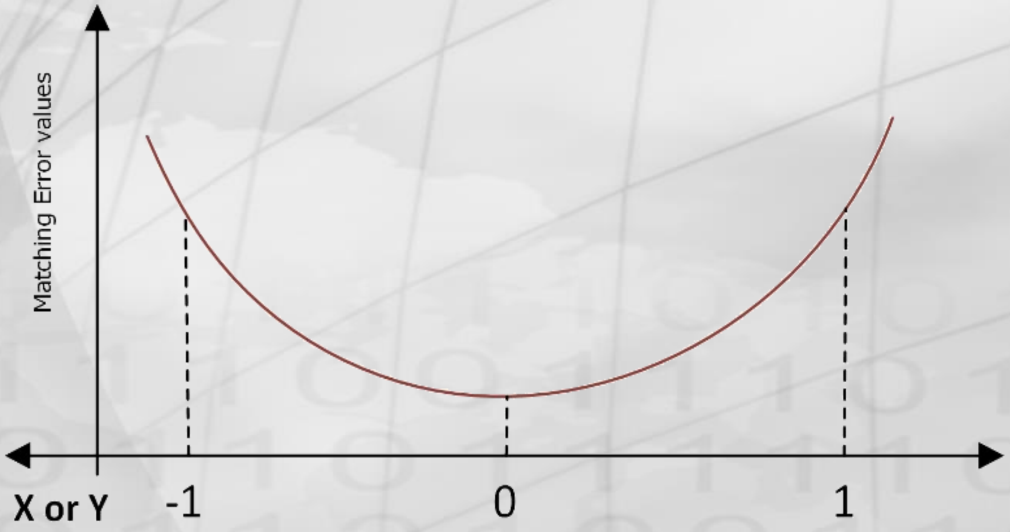
$$f_6(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$$

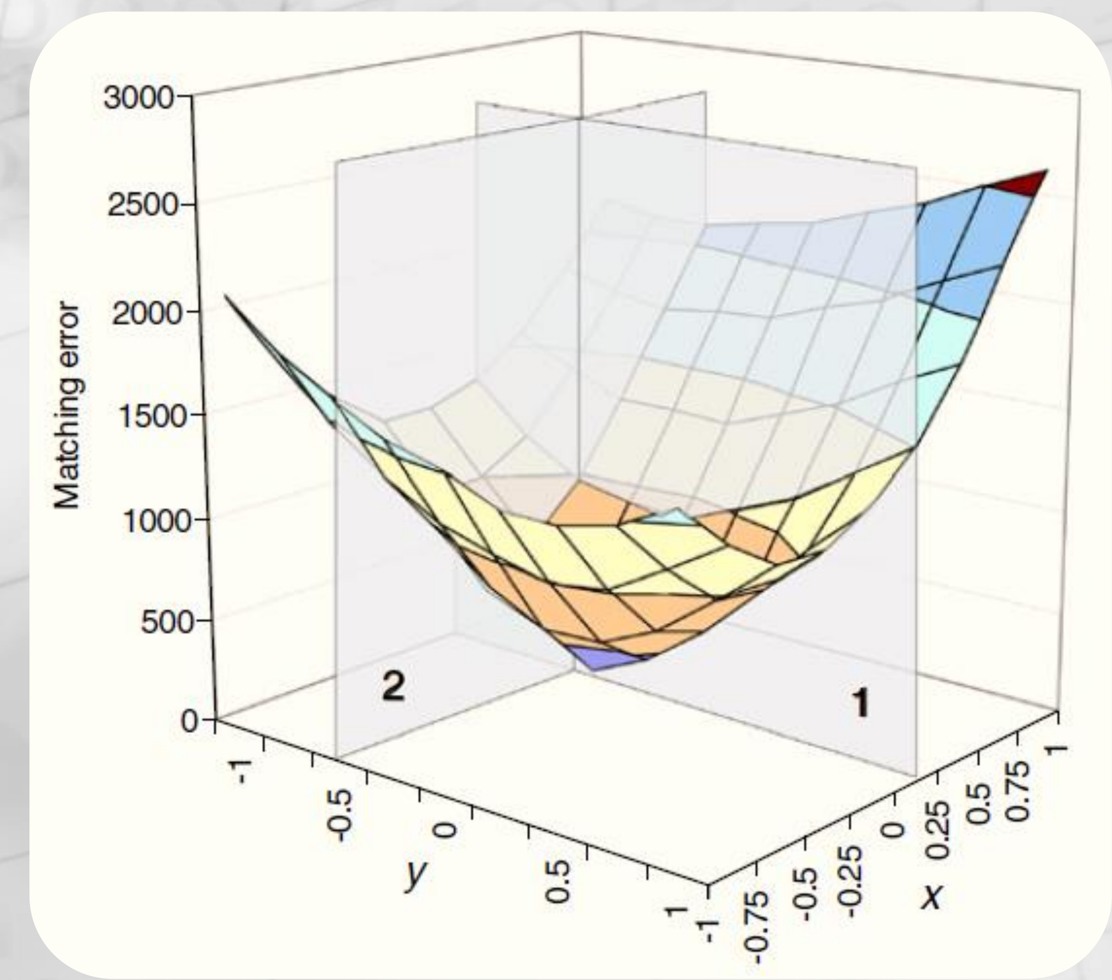
$$f_5(x,y) = Ax^2 + Bx + Cy^2 + Dy + E$$

Interpolation-Free Pixel Motion Estimation

- ***Sayed et al.*** [5] proposed to decompose the 2-D model of the error surface into 1-D parabolic curves, where any cross section with constant x or y in the error surface can be modeled with 1-D parabolic curve

$$f(t) = At^2 + Bt + C \qquad t = x \text{ or } y$$





Part IV: Proposed Algorithm.

- Mathematical Model.
- Algorithm Steps.
- Computational Complexity Analysis.
- Results

Proposed Algorithm

Mathematical Model

- 1-D error curve extended to higher order polynomial model in order to achieve higher prediction accuracy.

- a total of 25 cost values (5 rows by 5 columns including the best one at the origin P (0, 0)) are used to estimate the best fractional pixel location for each Prediction Unit (PU).

P(-2,-2)	P(-1,-2)	P(0,-2)	P(1,-2)	P(2,-2)
P(-2,-1)	P(-1,-1)	P(0,-1)	P(1,-1)	P(2,-1)
P(-2,0)	P(-1,0)	P(0,0)	P(1,0)	P(2,0)
P(-2,1)	P(-1,1)	P(0,1)	P(1,1)	P(2,1)
P(-2,2)	P(-1,2)	P(0,2)	P(1,2)	P(2,2)

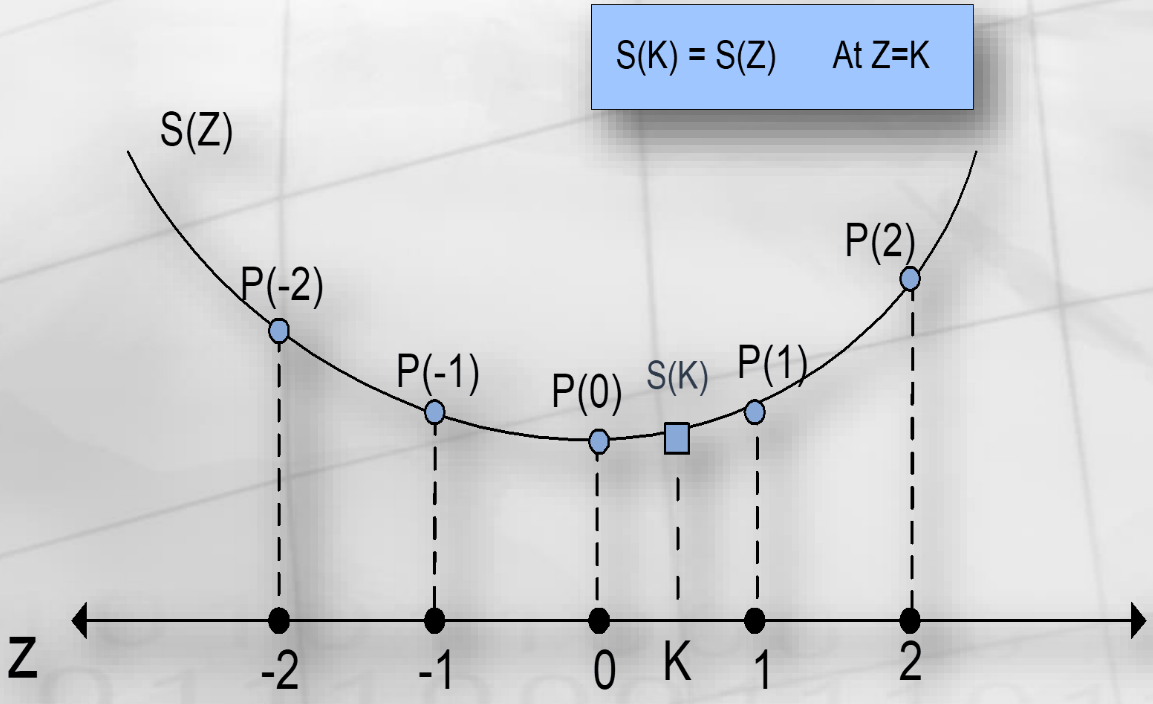
Proposed Algorithm

- The 1-D curve can be estimated using mathematical interpolation. Our algorithm applies Lagrange interpolation for the estimation process.

$$S(Z) = \sum_{i=-2}^{2} \mathrm{P}(i)\, L_i(z) \qquad Z = x \text{ or } y$$

$$L_i(Z) = \prod_{j=0,j\neq i}^{n} \frac{Z-Z_i}{Z_i-Z_j} = \frac{(Z-Z_0)\ldots\ldots(Z-Z_{i-1})(Z-Z_{i+1})\ldots(Z-Z_n)}{(Z_i-Z_0)\ldots\ldots(Z_i-Z_{i-1})(Z_i-Z_{i+1})\ldots\ldots(Z_i-Z_n)}$$

- where P(i) is the matching error value at each location i in 1-D , i = [-2,-1, 0, 1, 2].



Proposed Algorithm

- By substitution the resulted curve model is a fourth degree polynomial equation as described in following Equation:

$$\boldsymbol{S(Z) = C1{*}Z^4 + C2{*}Z^3 + C3{*}Z^2 + C4{*}Z + C5.}$$

- The five constants *C1* to *C5* can be calculated using the known five matching error values as in Equation:

$$\begin{bmatrix} C1 \\ C2 \\ C3 \\ C4 \\ C5 \end{bmatrix} = \begin{bmatrix} \frac{1}{24} & -\frac{1}{6} & \frac{1}{4} & -\frac{1}{6} & \frac{1}{24} \\ -\frac{1}{12} & \frac{1}{6} & 0 & -\frac{1}{6} & \frac{1}{12} \\ -\frac{1}{24} & \frac{4}{6} & -\frac{5}{4} & \frac{4}{6} & -\frac{1}{24} \\ \frac{1}{12} & -\frac{4}{6} & 0 & \frac{4}{6} & -\frac{1}{12} \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \mathrm{X} \begin{bmatrix} P(-2) \\ P(-1) \\ P(0) \\ P(1) \\ P(2) \end{bmatrix}$$

Proposed Algorithm

- The minimum error location is most likely to fall within the range -0.5 ≤ x, y ≤ 0.5. So [5], the proposed algorithm estimates the matching error at the locations within these ranges.

- The matching error values at fractional pixel locations (-0.5, -0.25, 0, 0.25, 0.5) can be estimated using the calculated coefficients as:

$$\begin{bmatrix} P(-0.5) \\ P(-0.25) \\ P(0.25) \\ P(0.5) \end{bmatrix} = \begin{bmatrix} \frac{1}{16} & -\frac{1}{8} & \frac{1}{4} & -\frac{1}{2} & 1 \\ \frac{1}{256} & -\frac{1}{64} & \frac{1}{16} & -\frac{1}{4} & 1 \\ \frac{1}{256} & \frac{1}{64} & \frac{1}{16} & \frac{1}{4} & 1 \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & 1 \end{bmatrix} \text{x} \begin{bmatrix} C1 \\ C2 \\ C3 \\ C4 \\ C5 \end{bmatrix}$$

P(-0.5) **P(-0.25)** **P(0)** **P(0.25)** **P(0.5)**

Proposed Algorithm

- To reduce the algorithm complexity. We can observe that Z^4 and Z^3 have very small values for the fractional pixel locations (i.e. ±0.5 and ±0.25) since Z represents x or y. Therefore, the values of these terms have negligible contribution in the matching error estimation at the fractional pixel locations. The simplified Equations will become:

$$\begin{bmatrix} C3 \\ C4 \\ C5 \end{bmatrix} = \begin{bmatrix} -\frac{1}{24} & \frac{4}{6} & -\frac{5}{4} & \frac{4}{6} & -\frac{1}{24} \\ \frac{1}{12} & -\frac{4}{6} & 0 & \frac{4}{6} & -\frac{1}{12} \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \mathrm{x} \begin{bmatrix} P(-2) \\ P(-1) \\ P(0) \\ P(1) \\ P(2) \end{bmatrix}$$

Calculating the coefficients

Calculating the matching error values at fractional locations

$$\begin{bmatrix} P(-0.5) \\ P(-0.25) \\ P(0.25) \\ P(0.5) \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & -\frac{1}{2} & 1 \\ \frac{1}{16} & -\frac{1}{4} & 1 \\ \frac{1}{16} & \frac{1}{4} & 1 \\ \frac{1}{4} & \frac{1}{2} & 1 \end{bmatrix} \mathrm{x} \begin{bmatrix} C3 \\ C4 \\ C5 \end{bmatrix}$$

Proposed Algorithm

E-JUST

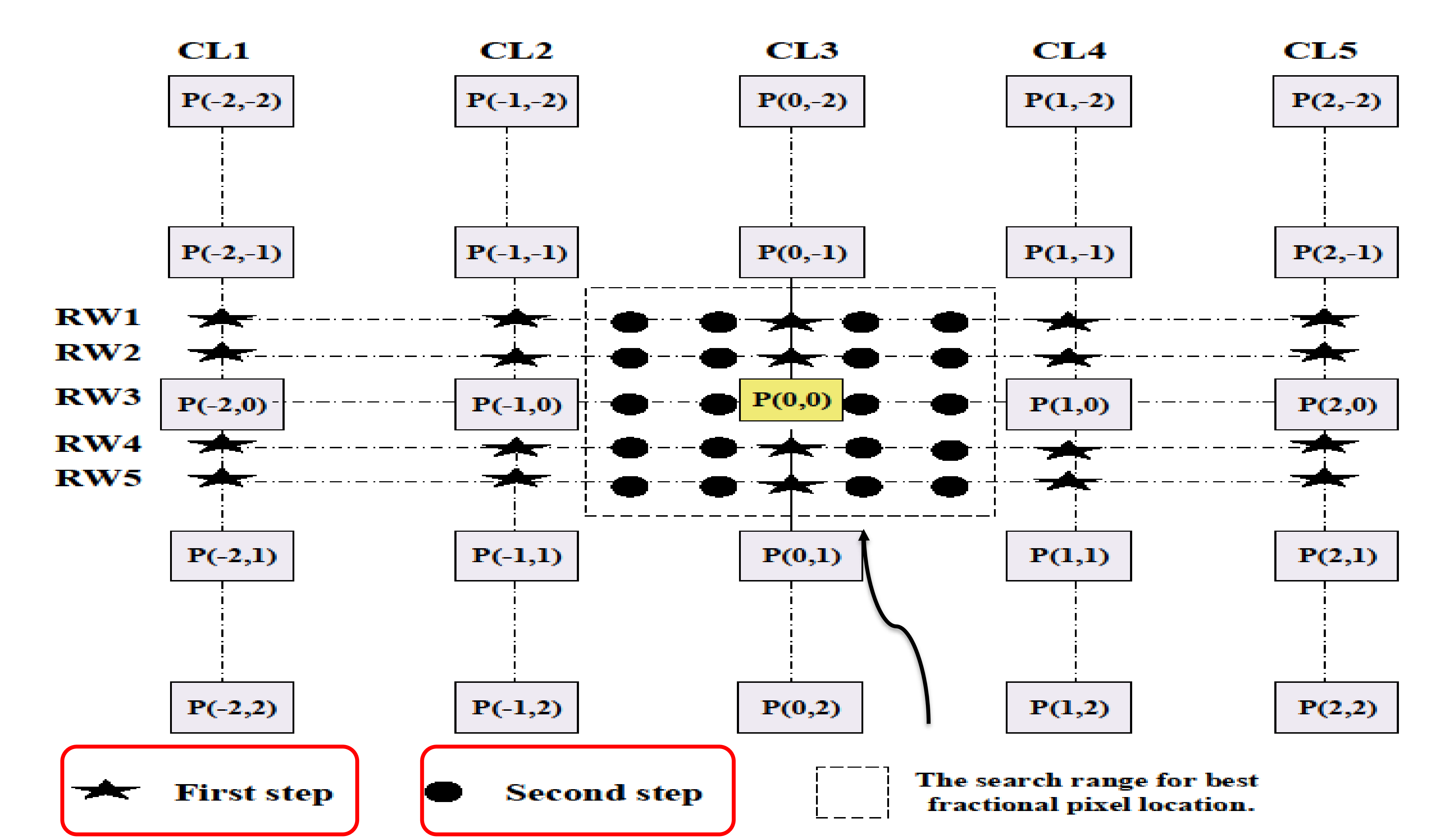
Algorithm Steps

1. Estimate the matching error at the fractional pixel locations represented with stars in the figure, using the 5 matching error values in each column from CL1 to CL5 using the previous equations vertically.

2. Apply the same equations horizontally at each row from RW1 to RW5 to find the matching error at the fractional-pixel locations (represented with circles)by using the matching error of the fractional pixel locations calculated in the previous step.

3. Find the location with the minimum matching error value among the 25 fractional pixel locations. This location is the best fractional pixel location.

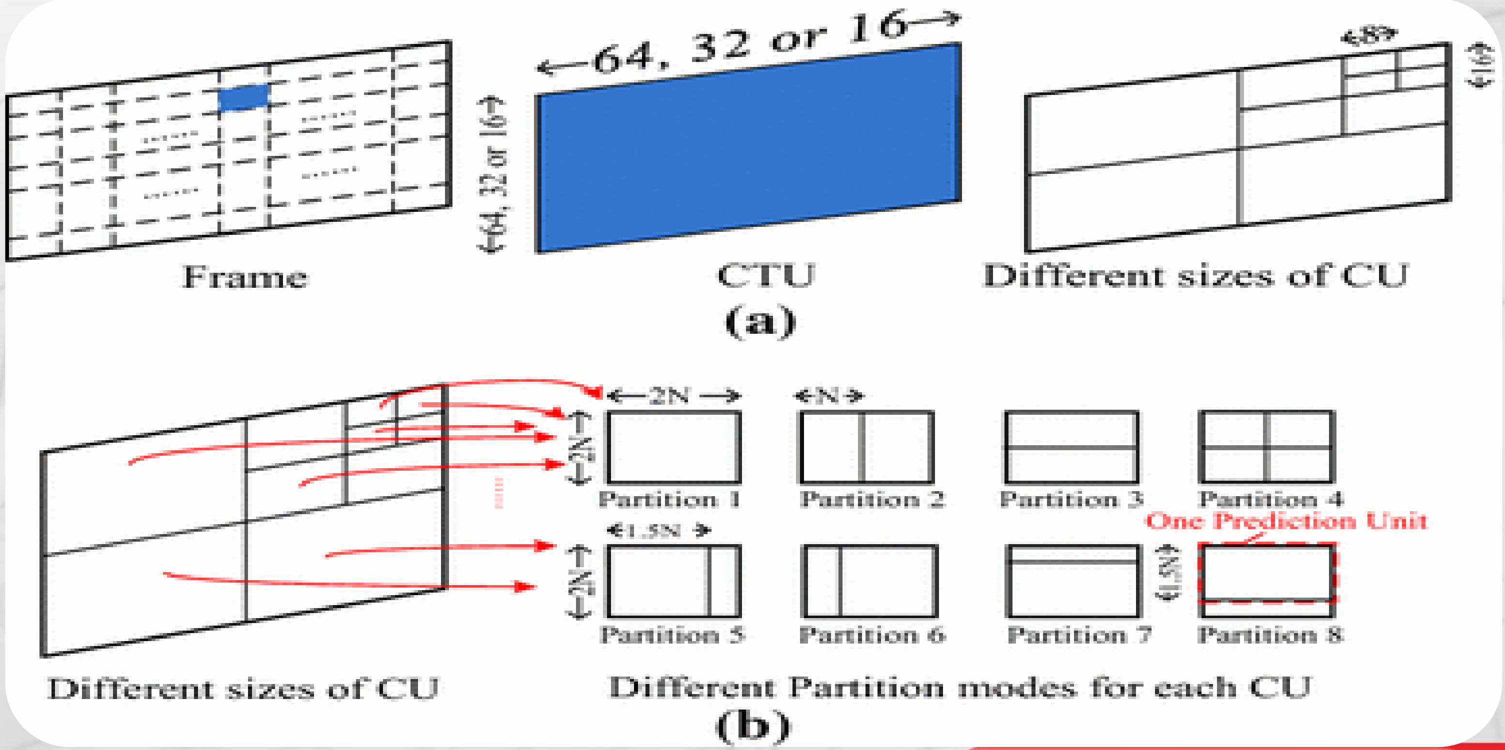
Proposed Algorithm



Proposed Algorithm

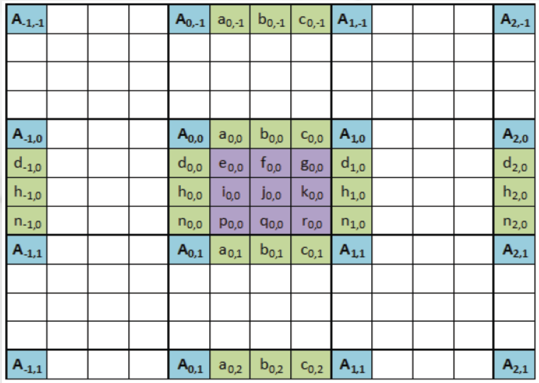
Computational Complexity Analysis

- HEVC adopted 24 Prediction Unit (PU) sizes [1] ranged from 4X8 and 8X4 to 64x64. All various PU partitioning combinations are examined by the ME encoder unit in order to decide which one of them gives the best results in terms of rate and distortion.



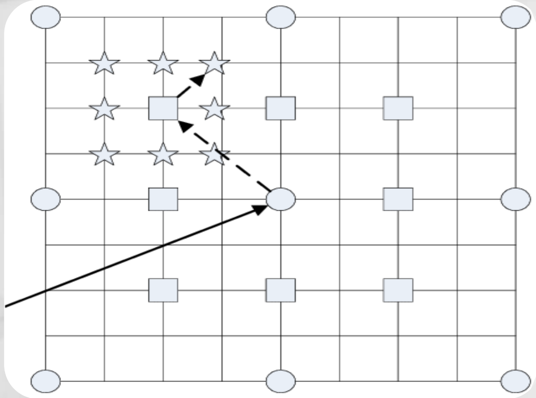
Proposed Algorithm

- The computational analysis of the proposed algorithm compared to traditional scheme (Hierarchal method in HM).



Traditional HM scheme

1-The FME unit interpolate the sub-pixel locations using the integer pixel of the reference frame for each Prediction Block (PB). For example, 16x16 block requires 54145 add/sub as in [9].



2-The FME unit calculates the cost of 16 fractional locations (8 for half , 8 for quarter), it requires N subtract and N-1 add operations for each block, where N is the number of pixels in each block

Proposed Algorithm

E-JUST

The proposed algorithm

25 Integer location Matching error values

25 calculated fractional location Matching error values

IME unit

Mathematical model

Search among 25 calculated Matching error values to estimate the best one has a minimum value

Requires 17x10 at worst case add/sub operation for each PB

Proposed Algorithm

- As a rough estimation :

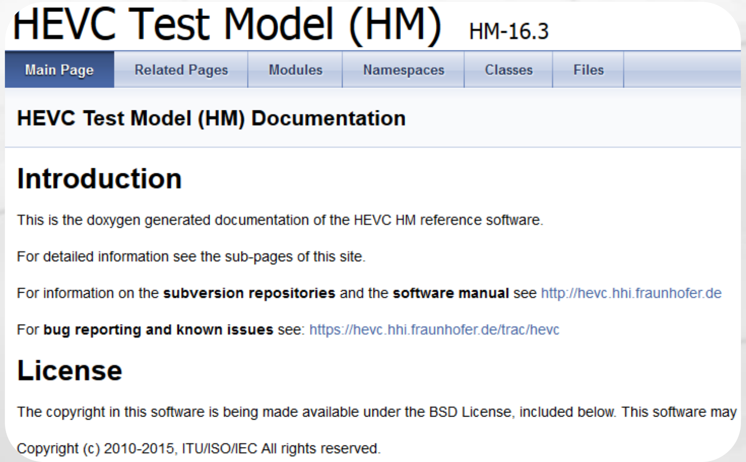
- **The proposed algorithm save up to 98% of the computational cost in comparison to traditional scheme for HM, achieved at the case of full search method in IME unit.**

- **At the case of fact search method in IME unit, the proposed algorithm saves computational cost h averaged about 86.6% (need extra step to extract 25 point from the IME unit).**

E-JUST

Results

- Algorithms in [5] and [6] were implemented in HEVC standard software HM-16.9.



- QP=22, 27, 32 and 37.
- search range: ±64.
- TZS as a fast search Integer Pixel Motion Estimation.
- All video sequences of class B, C, D, E, and F were used in the simulation.
- The matching error criterion is SSE (Sum of Squared Error).

Proposed Algorithm

	[5]		[6]		Proposed	
	BD-Bitrate (%)	BD-PSNR(dB)	BD-Bitrate (%)	BD-PSNR(dB)	BD-Bitrate (%)	BD-PSNR(dB)
Class B						
BQTerrace	4.4	-0.06	4.2	-0.05	1.3	-0.01
BasketballDrive	2.3	-0.05	1.8	-0.04	0.8	-0.02
Cactus	2.2	-0.04	2.3	-0.05	0.7	-0.01
Kimono	1.4	-0.04	0.7	-0.02	0.5	-0.01
ParkScene	2	-0.06	1.4	-0.04	0.5	-0.01
Class C						
BQMall	2.2	-0.08	2.9	-0.1	1.3	-0.05
BasketballDrill	1.4	-0.05	2.4	-0.09	0.9	-0.03
PartyScene	3.4	-0.12	4.4	-0.15	2	-0.07
RaceHorses	2.6	-0.09	3.6	-0.12	1.4	-0.05
Class D						
BQ Square	6	-0.18	8.5	-0.25	3.4	-0.1
BasketballPass	2.2	-0.1	3.1	-0.16	1.4	-0.06
BlowingBubbles	3.6	-0.12	4.6	-0.14	2.1	-0.07
RaceHorses	3.2	-0.13	4.5	-0.18	1.5	-0.06
Class E						
FourPeople	1.1	-0.03	1.6	-0.05	0.9	-0.02
Johnny	1.8	-0.04	2.5	-0.05	0.7	-0.02
KristenAndSara	1.2	-0.03	1.6	-0.04	0.7	-0.02
Class F						
BasketballDrillText	1.5	-0.06	2.8	-0.1	1	-0.04
ChinaSpeed	1.2	-0.05	5	-0.23	0.8	-0.04
SlideEditing	2.3	-0.3	3.1	-0.4	1.6	-0.2
SlideShow	3.9	-0.27	6.4	-0.4	2.7	-0.2

Proposed Algorithm

Summary	Class B	2.4	-0.05	2	-0.04	0.7	-0.01
	Class C	2.4	-0.08	3.3	-0.11	1.4	-0.05
	Class D	3.7	-0.13	5.1	-0.18	2.1	-0.07
	Class E	1.3	-0.03	1.9	-0.04	0.7	-0.02
	Class F	2.2	-0.17	4.3	-0.28	1.5	-0.12
Avg.		2.5	-0.09	3.3	-0.13	1.3	-0.05

- The proposed algorithm achieves significant encoding improvement for all video sequences with respect to the existed algorithms in terms of BD-Rate and BD-PSNR in comparison with existed algorithms.

Conclusion

- HEVC is a new video coding raises the need for efficient hardware architecture.

- The traditional fractional pixel motion estimation using interpolation method suffers from high computational complexity, memory requirements and large encoding time.

- Interpolation- free methods have been proposed to limit the drawbacks of the traditional scheme , the performance degradation has presented.

- A new proposed interpolation-free algorithm is presented with higher computational saving and better performance.

Questions

References

[1] Sullivan, Gary J., Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, ”Overview of the high efficiency video coding (HEVC) standard” IEEE Transactions on circuits and systems for video technology.,vol. 22, no. 12 , Dec 2012.

[2] Dikbas, S., Arici, T. and Altunbasak, Y. (2010). Fast Motion Estimation With Interpolation-Free Sub-Sample Accuracy. IEEE Transactions on Circuits and Systems for Video Technology, 20(7), pp.1047-1051.

[3] Hill, P., Chiew, T., Bull, D. and Canagarajah, C. (2006). Interpolation Free Subpixel Accuracy Motion Estimation. IEEE Transactions on Circuits and Systems for Video Technology, 16(12), pp.1519-1526.

[4] Chang, J. and Leou, J. (2006). A quadratic prediction based fractional pixel motion estimation algorithm for H.264. Journal of Visual Communication and Image Representation, 17(5), pp.1074-1089.

[5] M. Sayed,W. Badawy, and G. Jullien, ”Interpolation-free fractional-pixel motion estimation algorithms with efficient hardware implementation.” Journal of Signal Processing Systems ,vol.67, no. 2 ,2012.

[6] Li, Y., Liu, Z., Ji, X. and Wang, D. (2016). HEVC fast FME algorithm using IME RD-costs based error surface fitting scheme. 2016 Visual Communications and Image Processing (VCIP), pp. 1-4.

[7] Lung, C.Y. and Shen, C.A., Design and implementation of a highly efficient fractional motion estimation for the HEVC encoder. Journal of Real-Time Image Processing, pp.1-17.

References

[8] V.Sze, M. Budagavi.(2014). Design and Implementa of next generation video coding systems (H.265/ HEVC tutorial),ISCAS,2014.

[9] Mert, Ahmet Can, Ercan Kalali, and Ilker Hamzaoglu. "Low complexity HEVC sub-pixel motion estimation technique and its hardware implementation." *Consumer Electronics-Berlin (ICCE-Berlin), 2016 IEEE 6th International Conference on*. IEEE, 2016.

THANKS FOR YOUR ATTENTION,,,,

By/Eng.Emad Badry
(emad.mahmoud@ejust.edu.eg)

The fifth IEEE Global Conference on Signal and Information Processing (GlobalSIP2017)