Universidad del Valle

Laboratorio de Multimedia y Vision

# A Comparison of Block-Matching Motion Estimation Algorithms

María Santamaría and María Trujillo

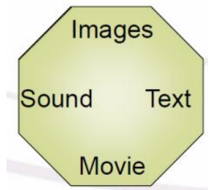

October 4th 2012
Séptimo Congreso Colombiano de Computación, 7CCC 2012, Medellín - Colombia

# Multimedia and Vision Laboratory

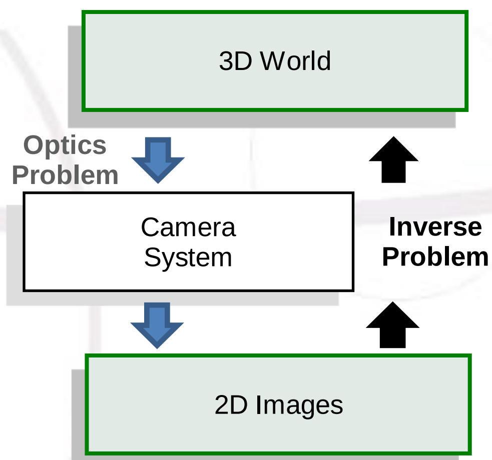
- MMV is a research group of the Universidad del Valle in Cali, Colombia

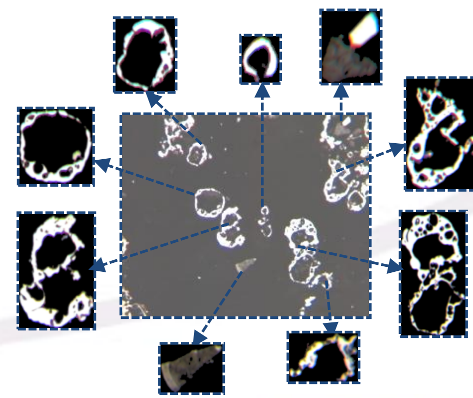
M. Santamaría

M. Trujillo




Multiple Media
&
Computer Vision



Multimedia and Vision Research Laboratory: http://mmv-lab.univalle.edu.co

Universidad del Valle

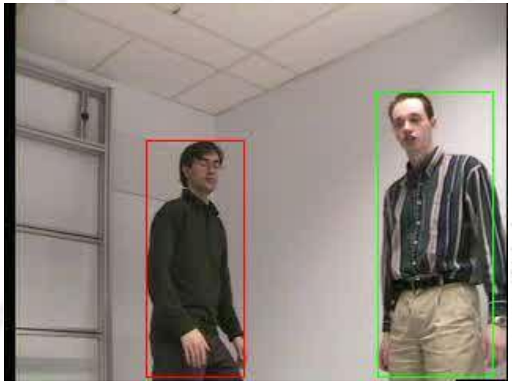
# Content

Laboratorio de Multimedia y Vision

- Motivation
- Motion Estimation
- Block-Matching
  - Distortion Metrics
  - Selected Algorithms
- Evaluation
  - Quality Metrics
  - Performance Metrics
- Video Test Sequences
- Results
- Final Remarks

# Motivation

Video coding



Tracking

3D TV

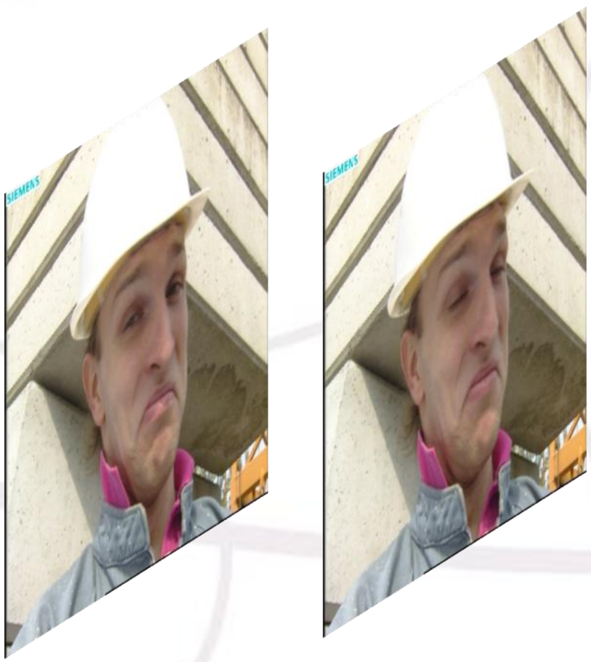
Gesture recognition

Resolution enhancement

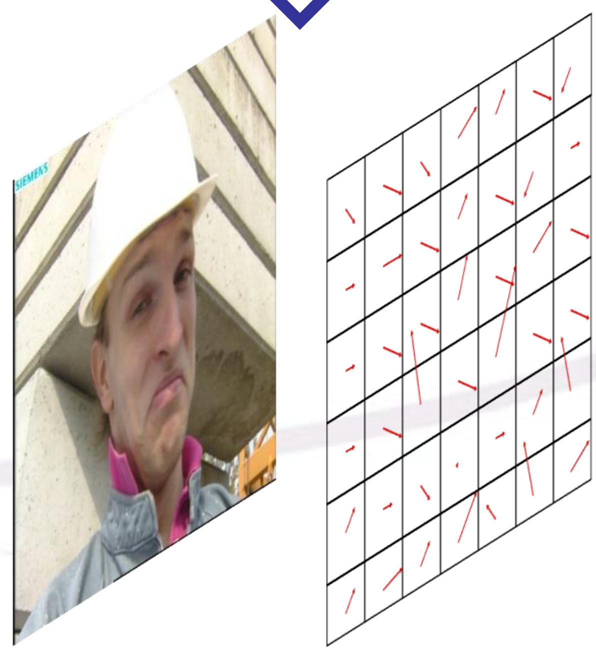
http://www.encodedmedia.com/
http://assets.vr-zone.net/15416/LGTV.jpg
http://csecar.wordpress.com/
http://www.newelectronics.co.uk/electronics-news/qualcomm-invests-in-gesture-recognition-technology/35620/
http://users.soe.ucsc.edu/~milanfar/research/resolution-enhancement.html

# Motion Estimation

Video Frames



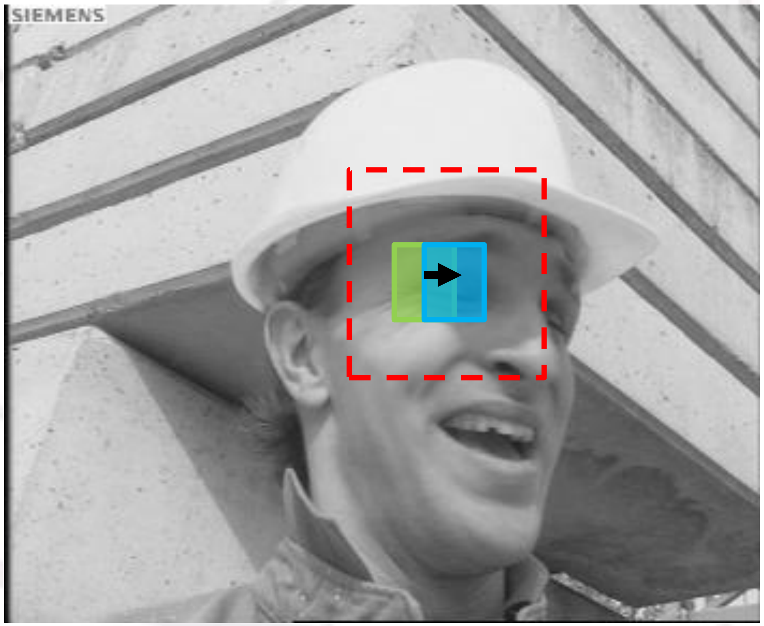
Motion Estimation



Motion Vectors

# Block-Matching

Reference Frame




Current Frame

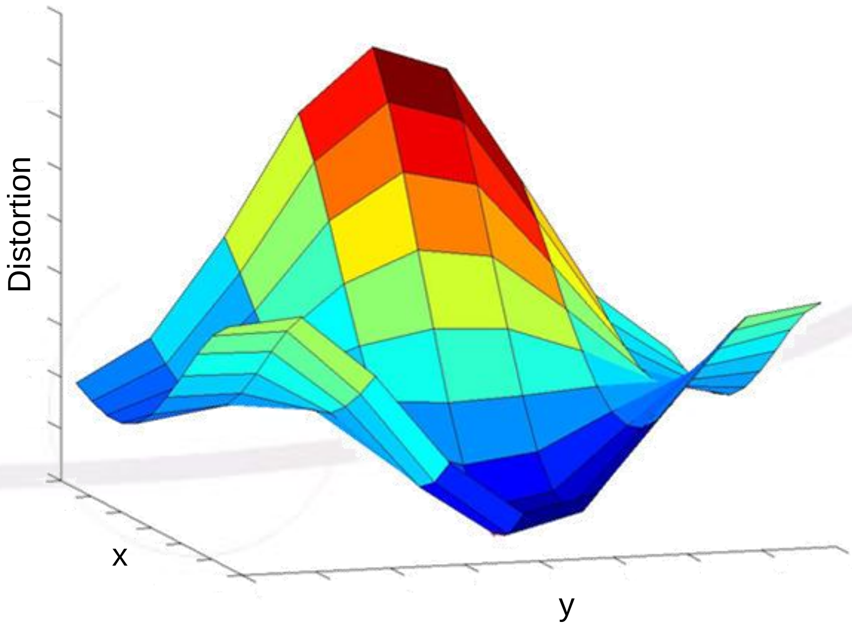

- Search area
- Current block
- Best matched block
- Motion vector

# Distortion Metrics

- The two most popular measures to determine the match between two blocks are: the Mean Square Error (MSE) and the Sum of Absolute Differences (SAD)

$$MSE(X,Y) = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}[X(i,j) - Y(i,j)]^2$$

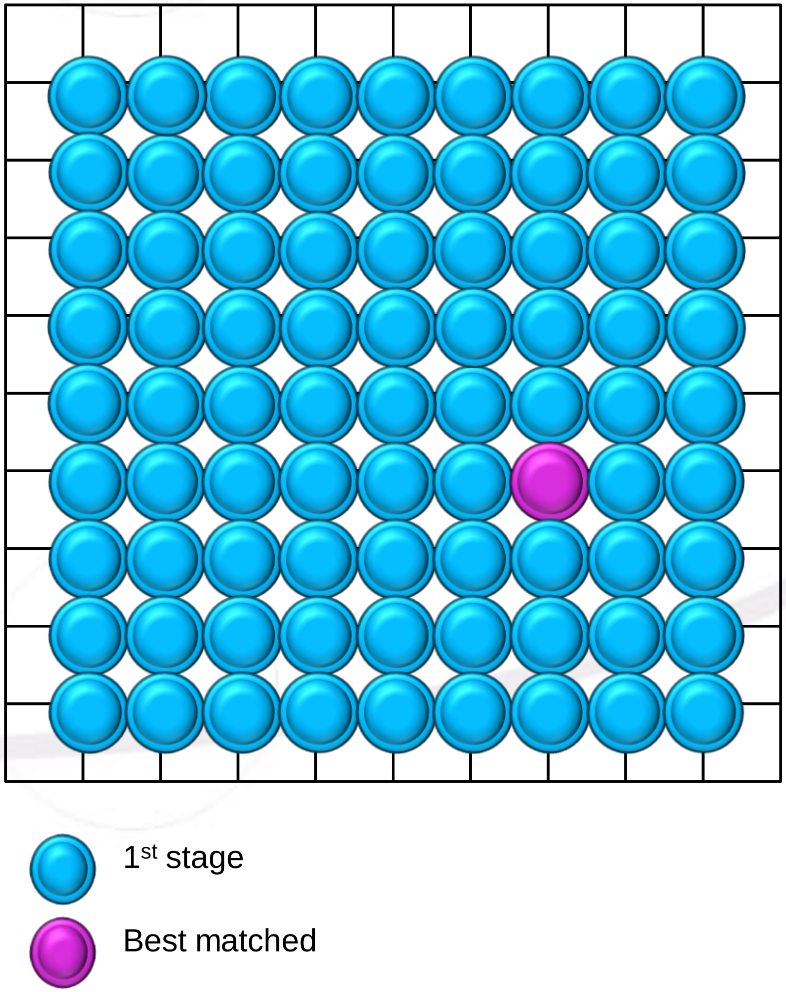
$$SAD(X,Y) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|X(i,j) - Y(i,j)|$$




B. Xiong and C. Zhu, "A new multiplication-free block matching criterion," IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 10, 2008
Elliot J. Rouse. A virtual curriculum vitae. http://www.elliottjrouse.com/

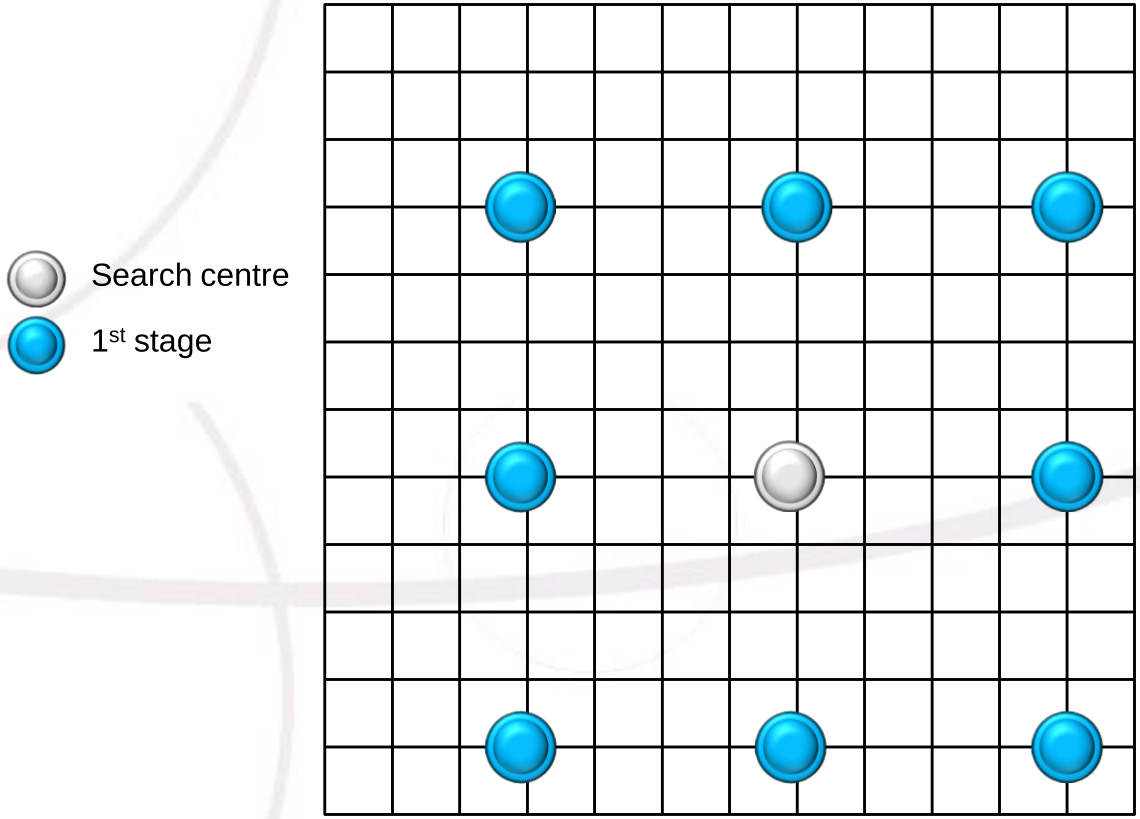
# Full-Search (FS)

- The Full-Search algorithm evaluates all positions in the window search of (2W+1) x (2W+1) size
- It involves high computational cost
- It is simple
- It guarantees a high accuracy in finding the best match



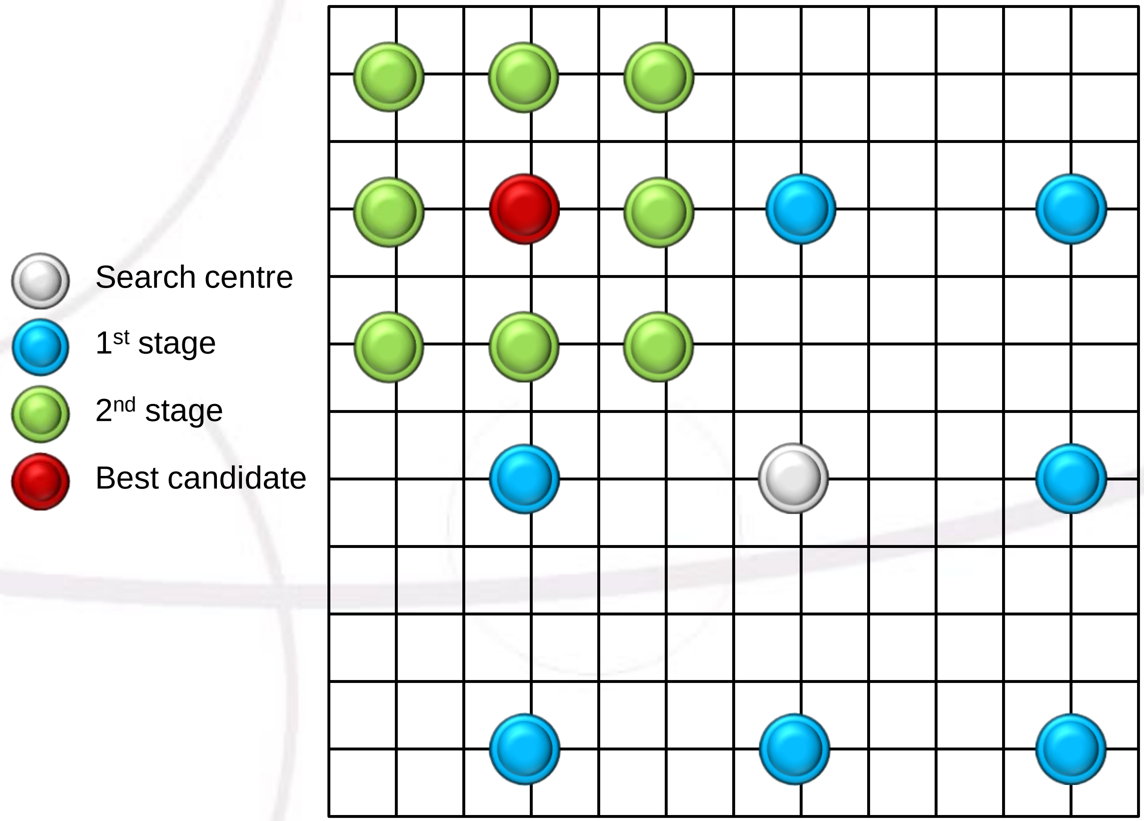

Y. Huzka, and P. Kulla, “Trends in Block-matching Motion Estimation Algorithms,” 2004

# Three-Step Search (3SS)



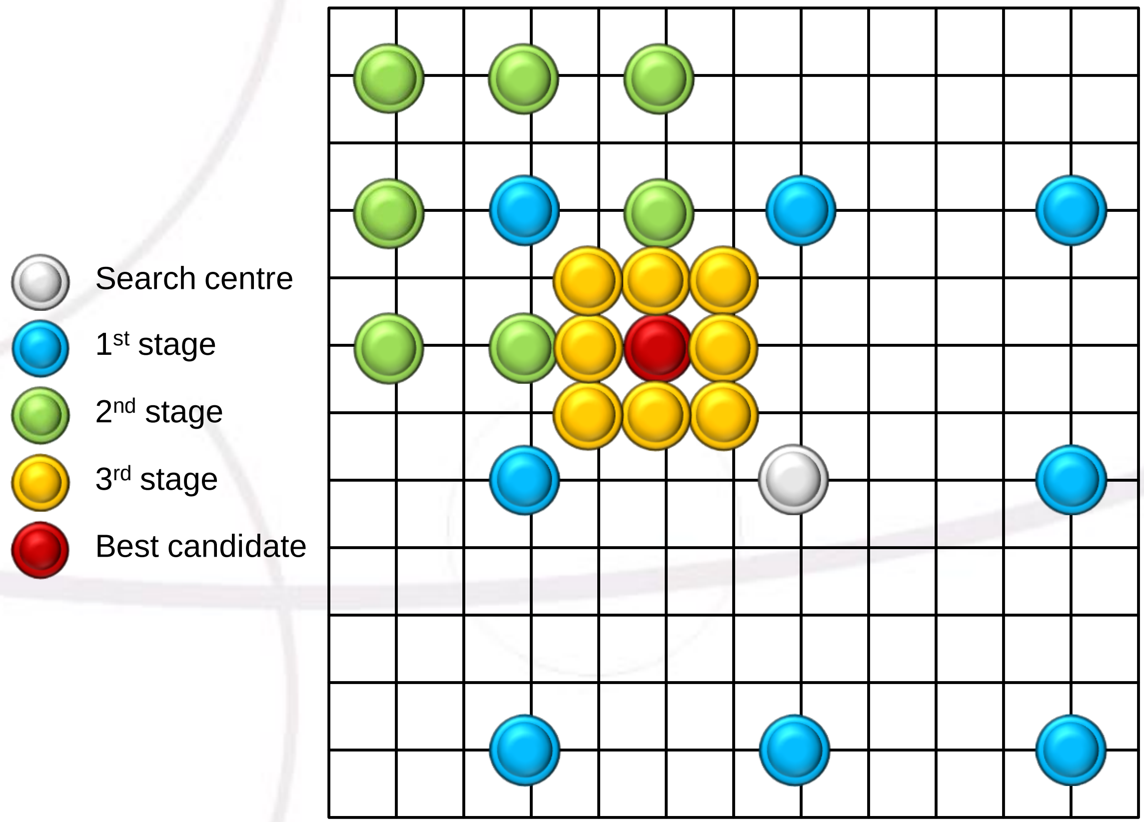

T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, “Motion Compensated Interframe Coding for Video Conferencing,” Proc. Nat. Telcommun. Conf., 1981

# Three-Step Search (3SS)




T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion Compensated Interframe Coding for Video Conferencing," Proc. Nat. Telcommun. Conf., 1981

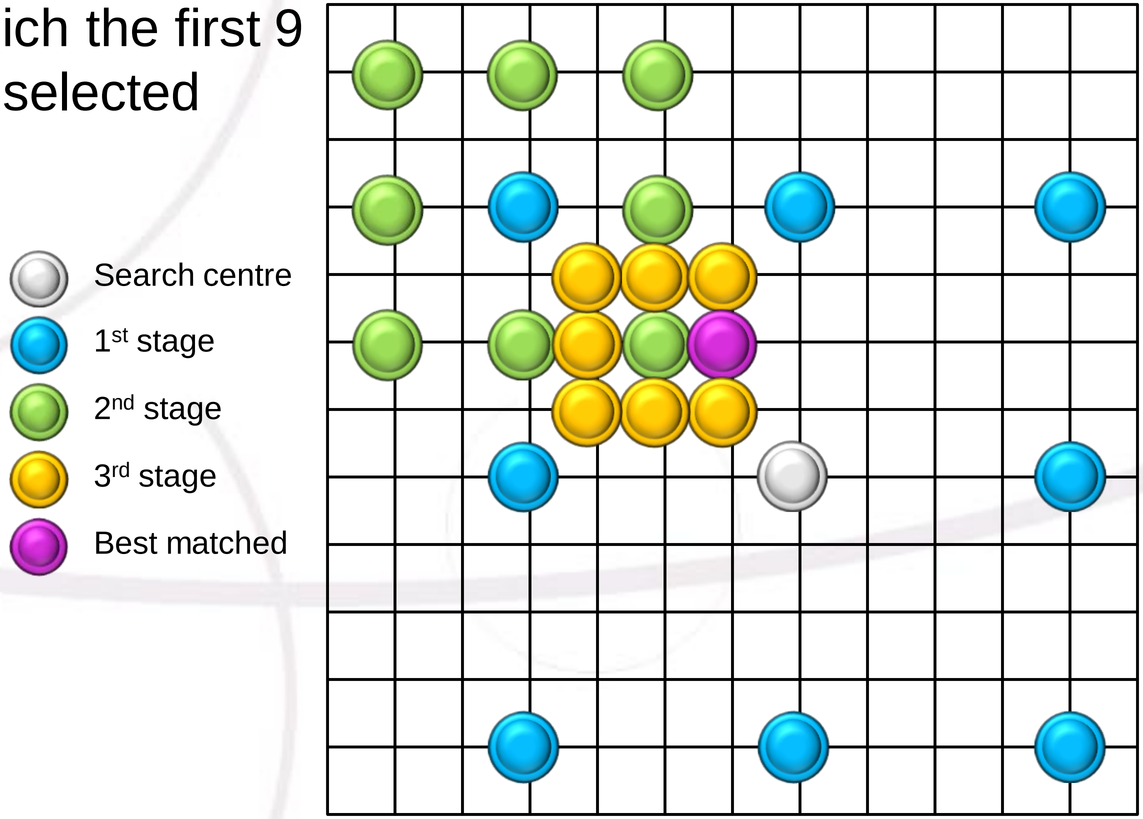
# Three-Step Search (3SS)




T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion Compensated Interframe Coding for Video Conferencing," Proc. Nat. Telcommun. Conf., 1981

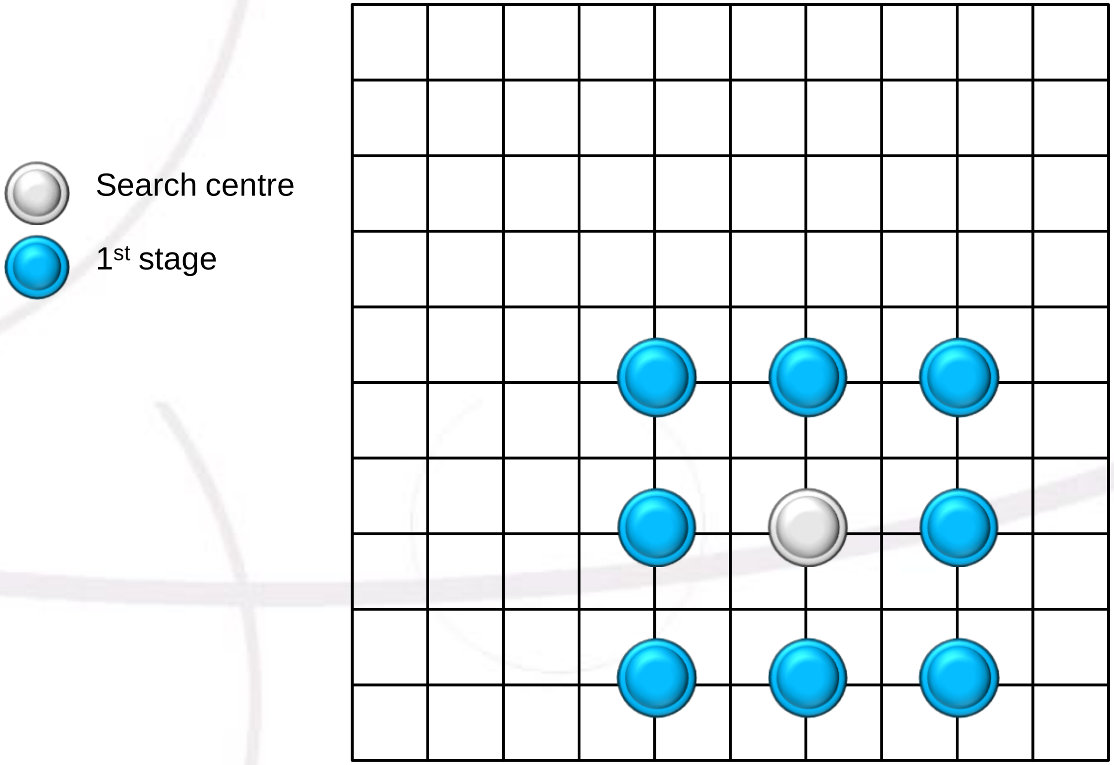
# Three-Step Search (3SS)

- The number of stages depends on the initial distance to which the first 9 neighbors are selected



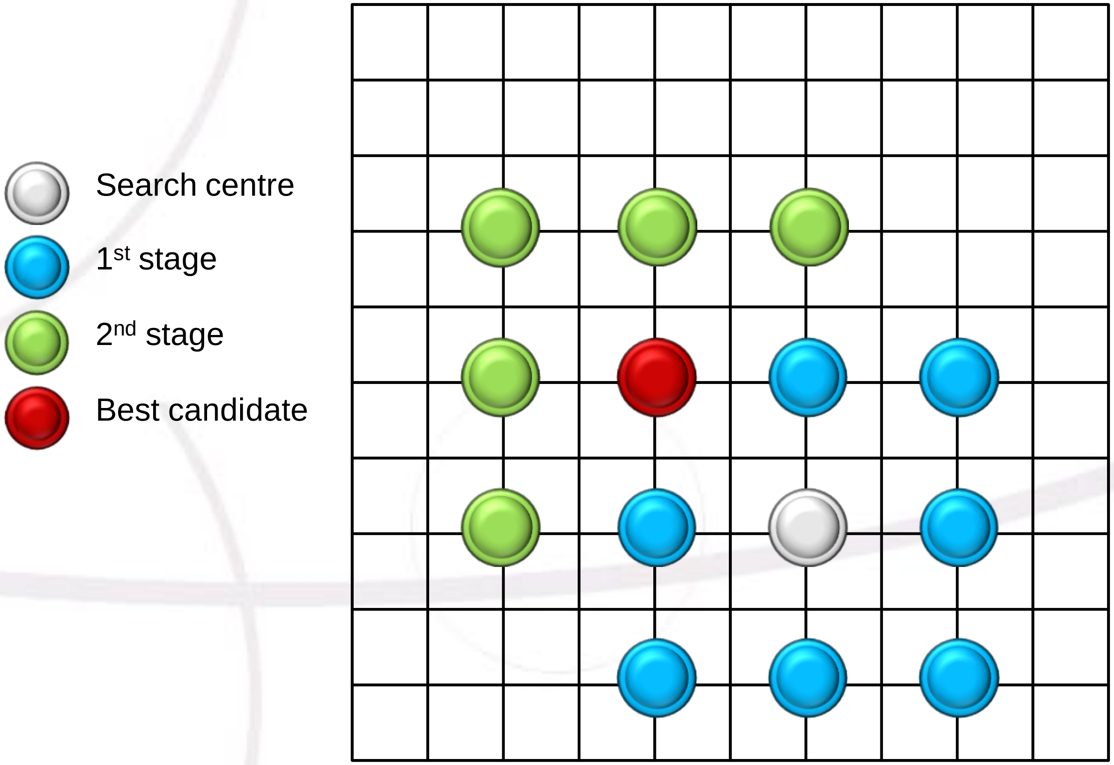

T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion Compensated Interframe Coding for Video Conferencing," Proc. Nat. Telcommun. Conf., 1981

# Four-Step Search (4SS)




L.-M. Po, and W. C.-Ma, "A novel four-step search algorithm for fast block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, 1996

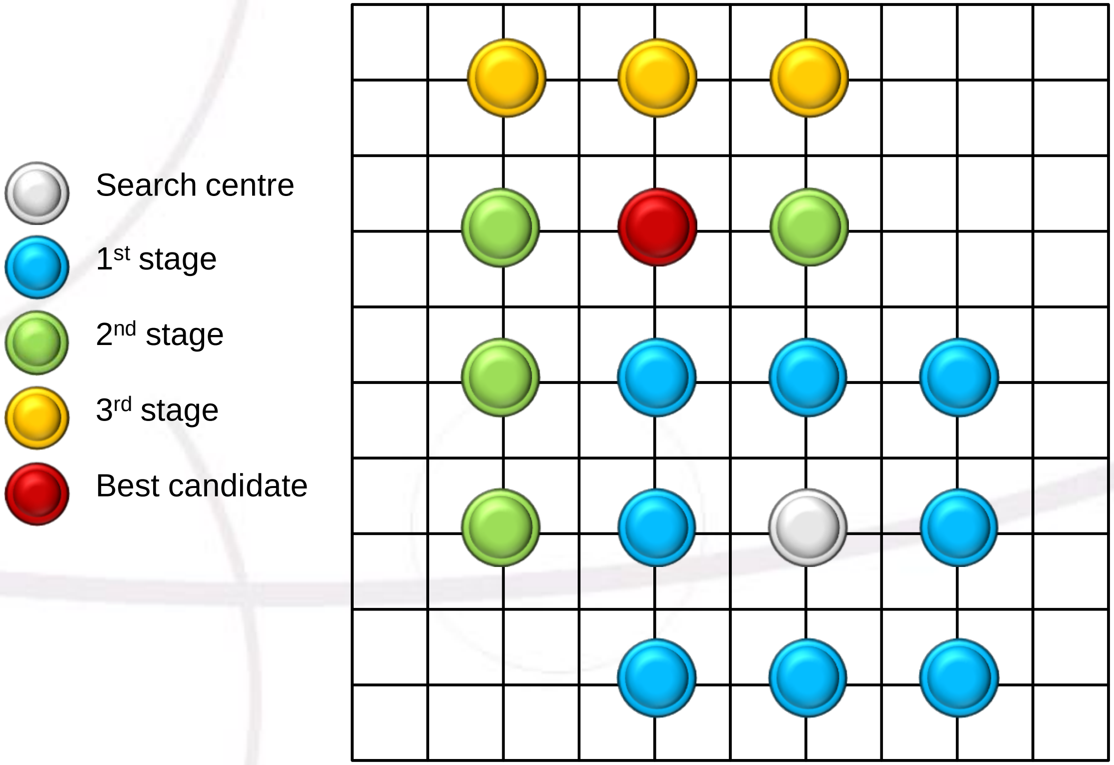
# Four-Step Search (4SS)




L.-M. Po, and W. C.-Ma, “A novel four-step search algorithm for fast block motion estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, 1996

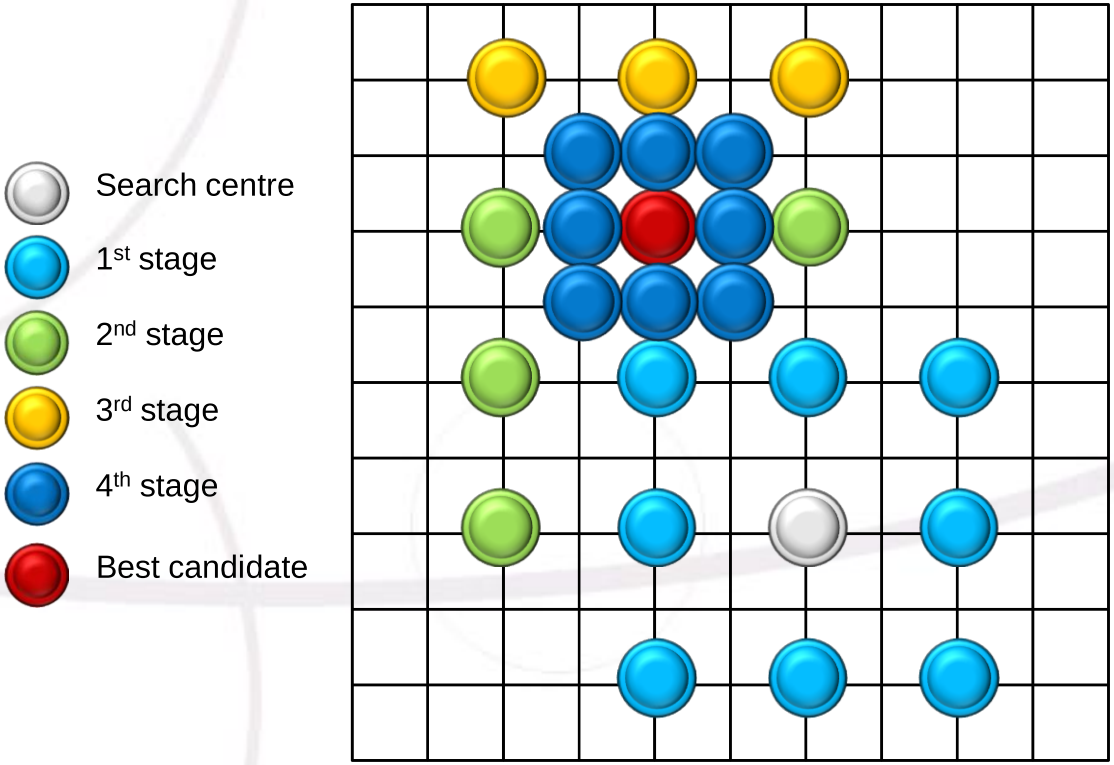
# Four-Step Search (4SS)






L.-M. Po, and W. C.-Ma, "A novel four-step search algorithm for fast block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, 1996

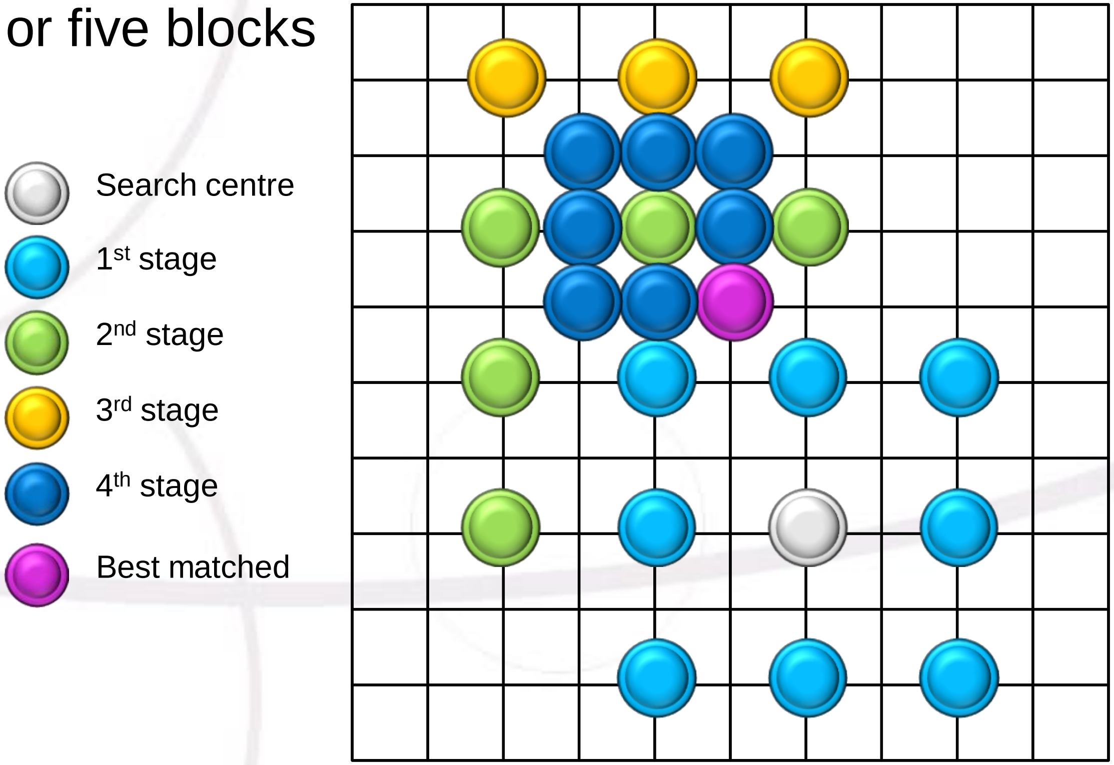
# Four-Step Search (4SS)




L.-M. Po, and W. C.-Ma, "A novel four-step search algorithm for fast block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, 1996

# Four-Step Search (4SS)

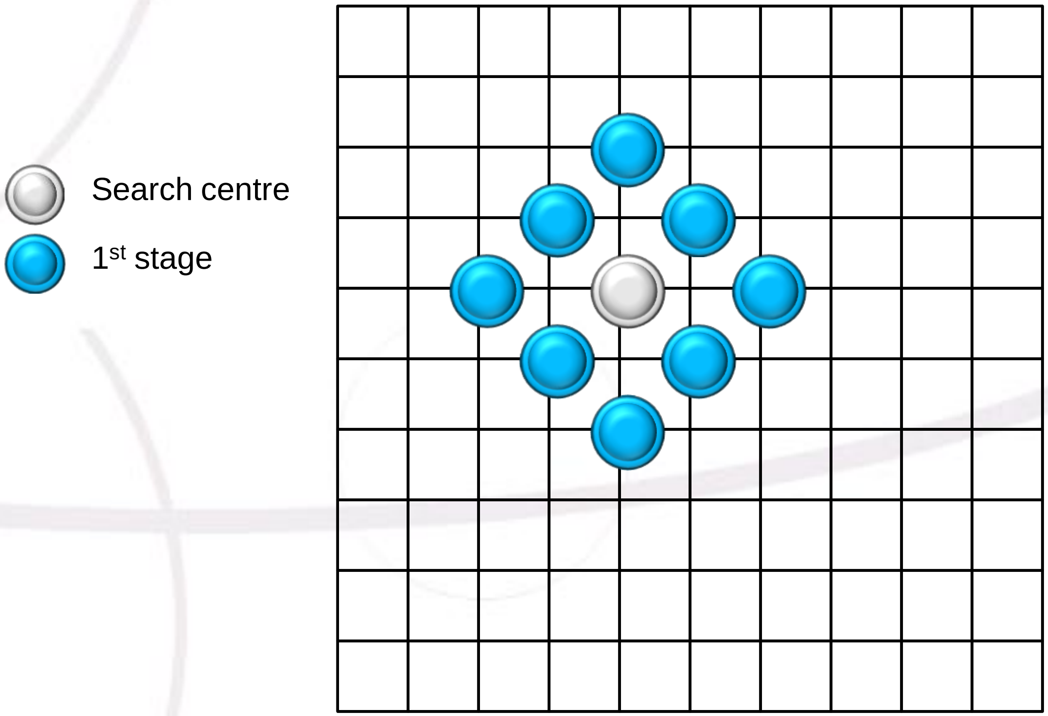
- Each new stage (except the reduced step stage) evaluates three or five blocks




L.-M. Po, and W. C.-Ma, “A novel four-step search algorithm for fast block motion estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, 1996

# Diamond Search (DS)

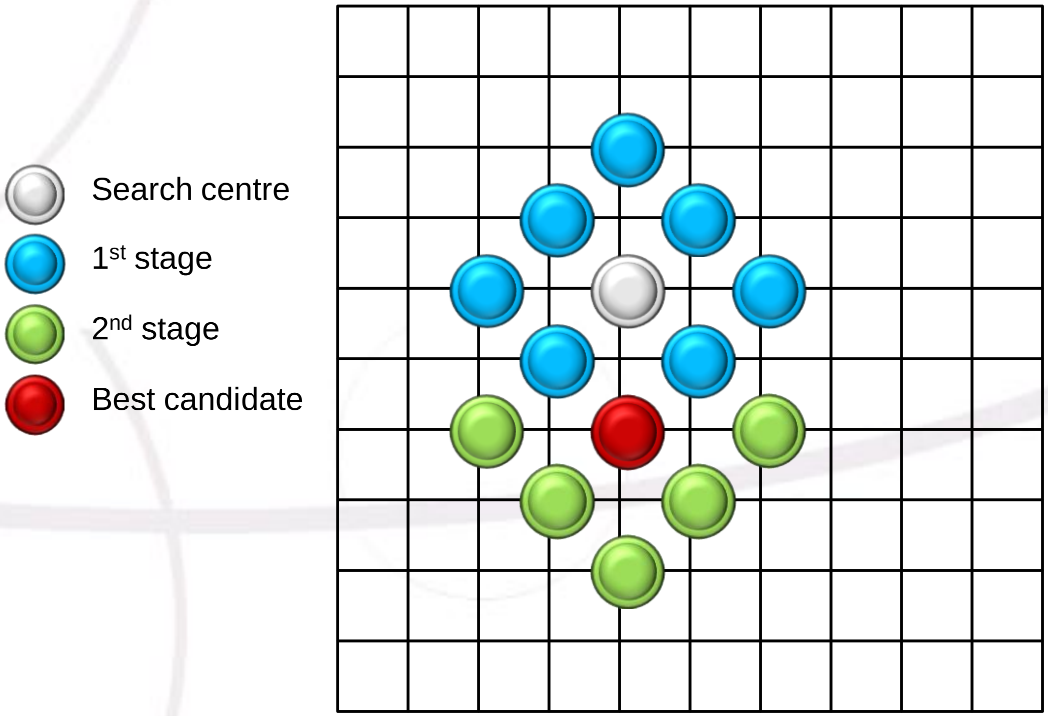





J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, “A novel unrestricted center-biased diamond search algorithm for block motion estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 4, 1998

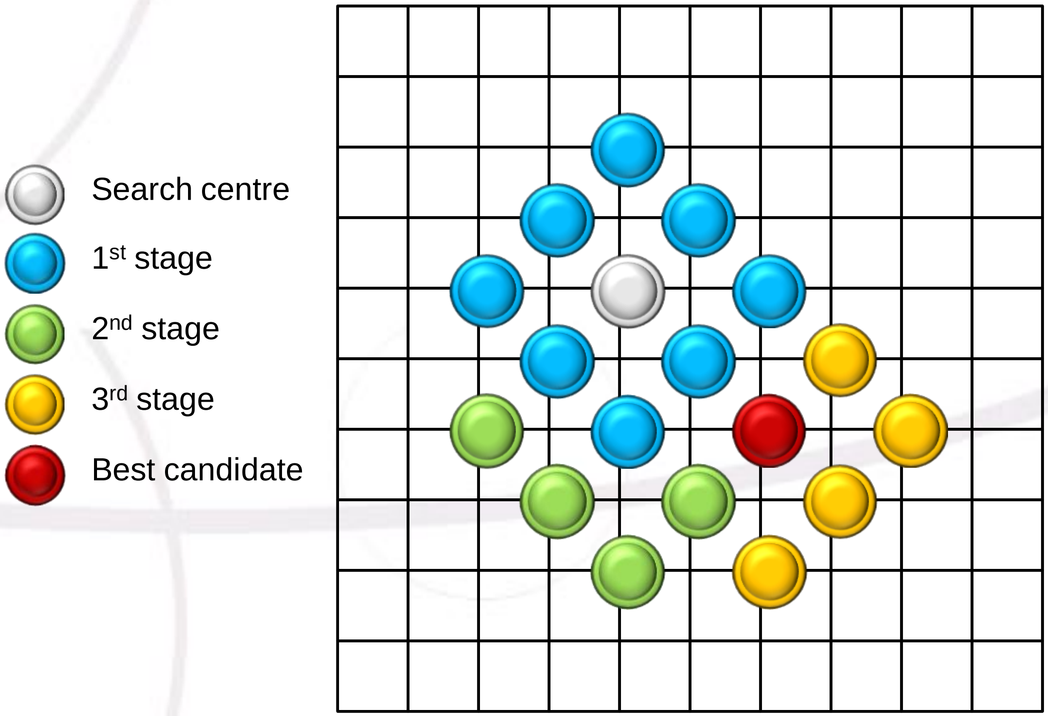
# Diamond Search (DS)






J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, “A novel unrestricted center-biased diamond search algorithm for block motion estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 4, 1998

# Diamond Search (DS)

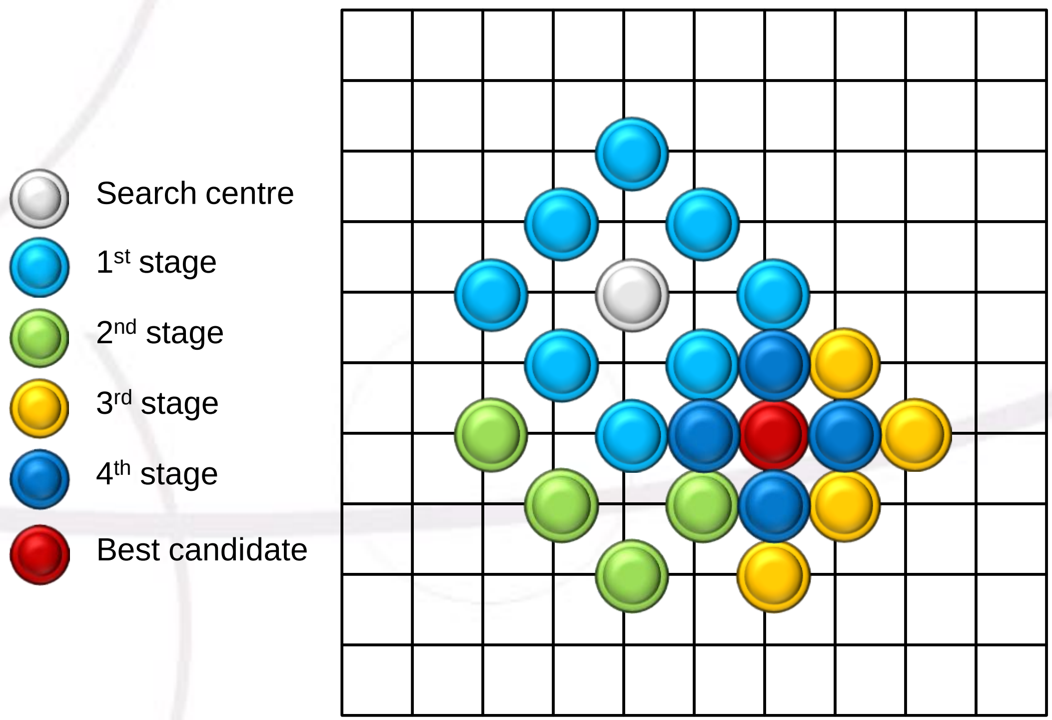



J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, “A novel unrestricted center-biased diamond search algorithm for block motion estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 4, 1998

Universidad del Valle

# Diamond Search (DS)

Laboratorio de Multimedia y Vision

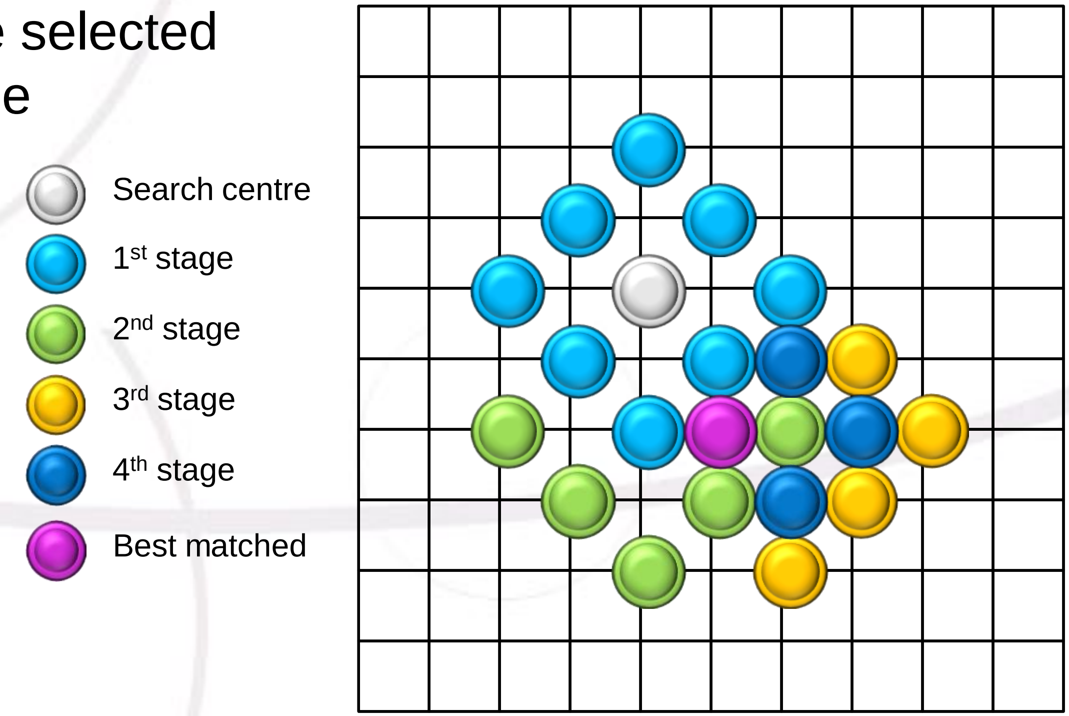



J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, “A novel unrestricted center-biased diamond search algorithm for block motion estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 4, 1998

# Diamond Search (DS)

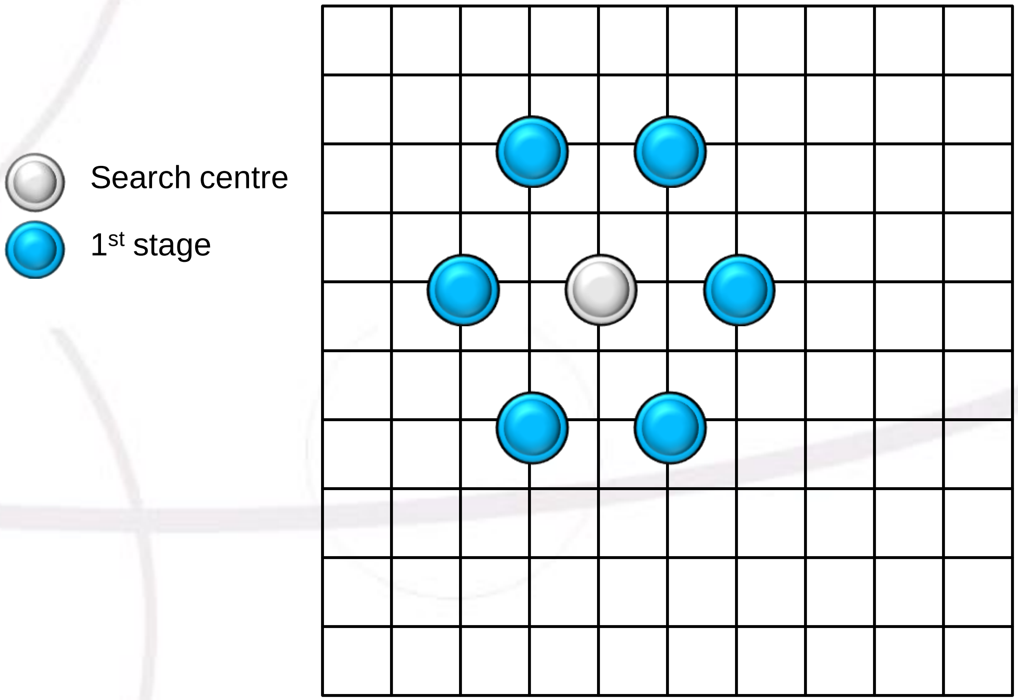


- Each new stage (except the reduced step stage) evaluates four or five blocks
- The neighbors are selected at a mixed distance



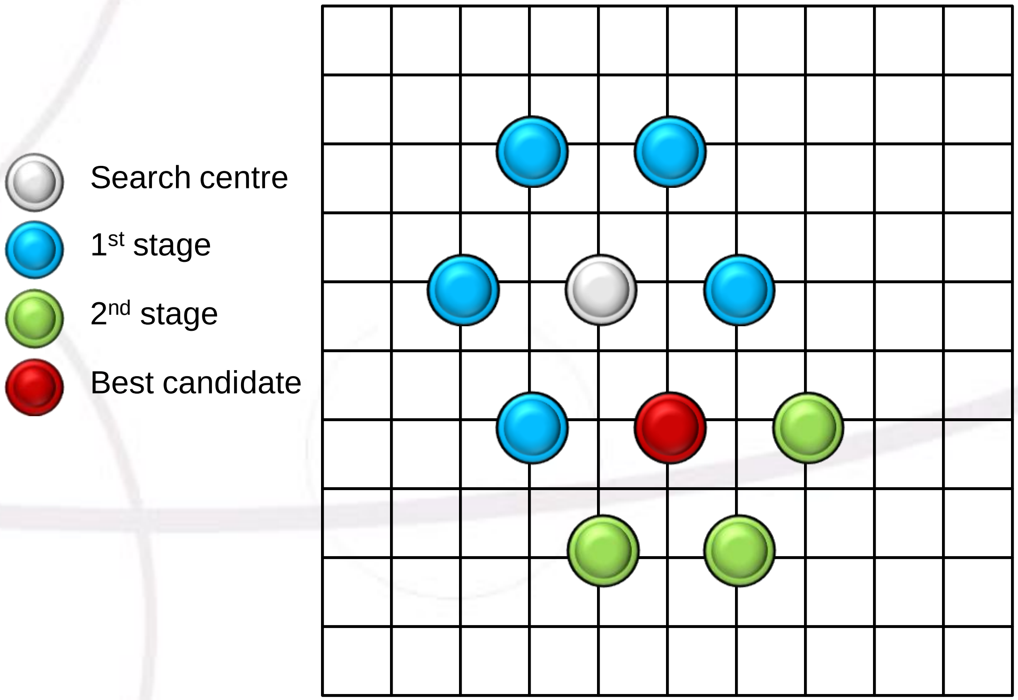

J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, "A novel unrestricted center-biased diamond search algorithm for block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 4, 1998

# Hexagonal Block Search (HEXBS)



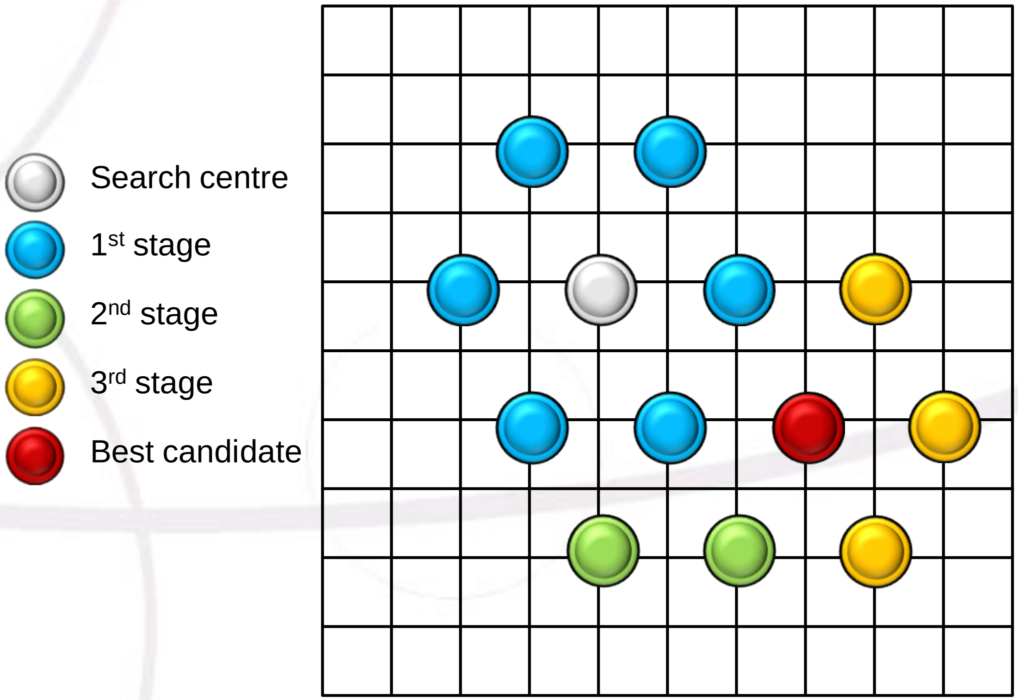

C.-H. Cheung and L.-M. Po, “Novel cross-diamond-hexagonal search algorithms for fast block motion estimation,” IEEE Trans. Multimedia, vol. 7, no. 1, 2005

# Hexagonal Block Search (HEXBS)




C.-H. Cheung and L.-M. Po, “Novel cross-diamond-hexagonal search algorithms for fast block motion estimation,” IEEE Trans. Multimedia, vol. 7, no. 1, 2005

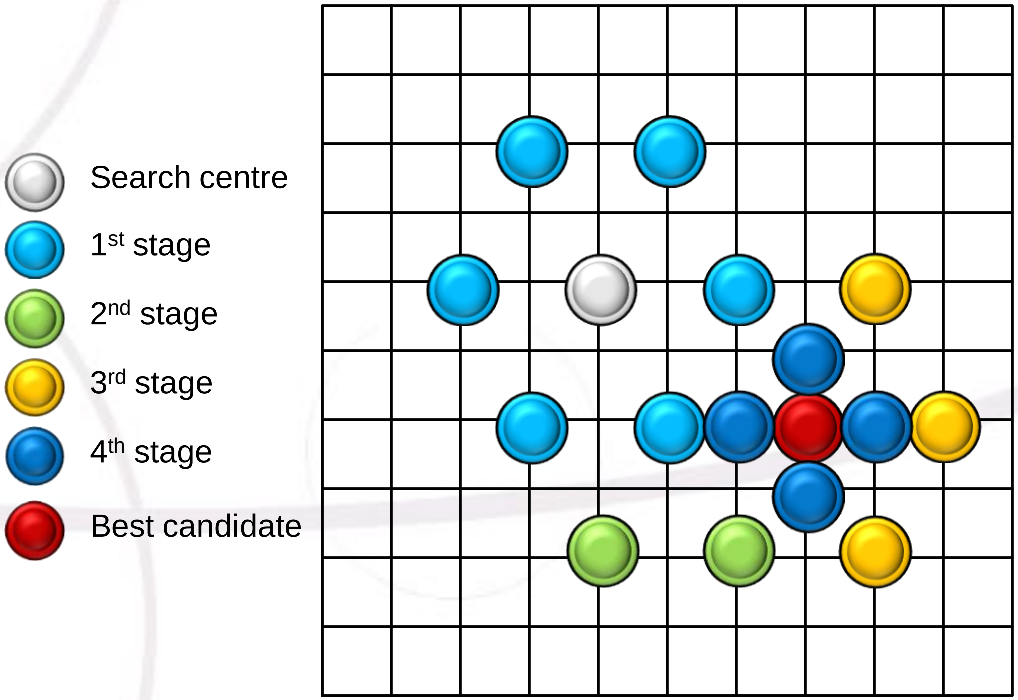
# Hexagonal Block Search (HEXBS)




C.-H. Cheung and L.-M. Po, “Novel cross-diamond-hexagonal search algorithms for fast block motion estimation,” IEEE Trans. Multimedia, vol. 7, no. 1, 2005

# Hexagonal Block Search (HEXBS)

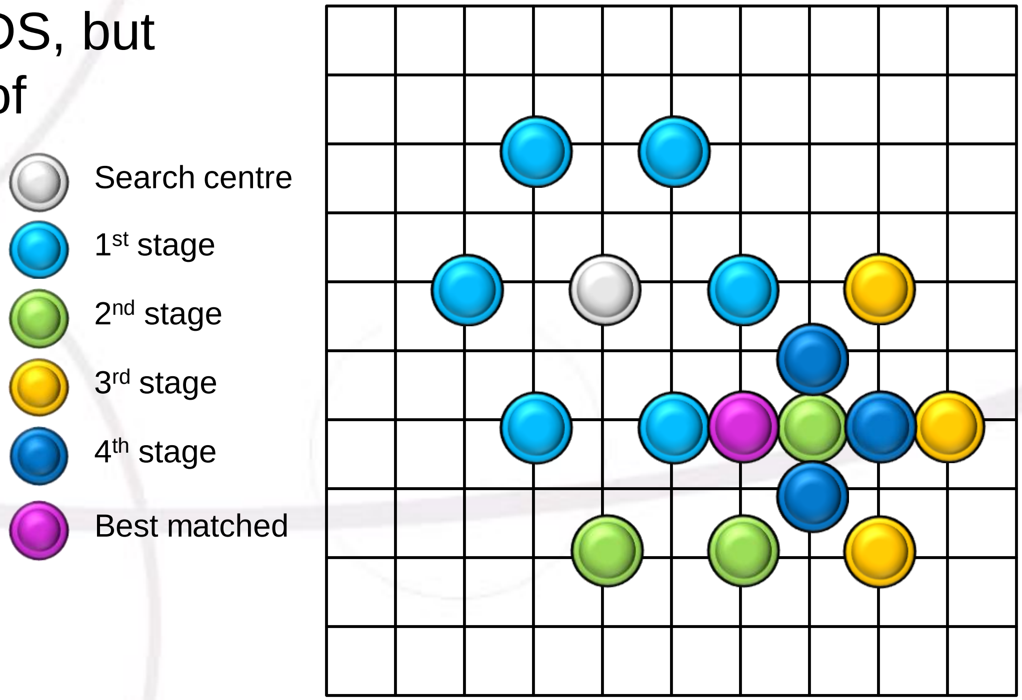





C.-H. Cheung and L.-M. Po, "Novel cross-diamond-hexagonal search algorithms for fast block motion estimation," IEEE Trans. Multimedia, vol. 7, no. 1, 2005

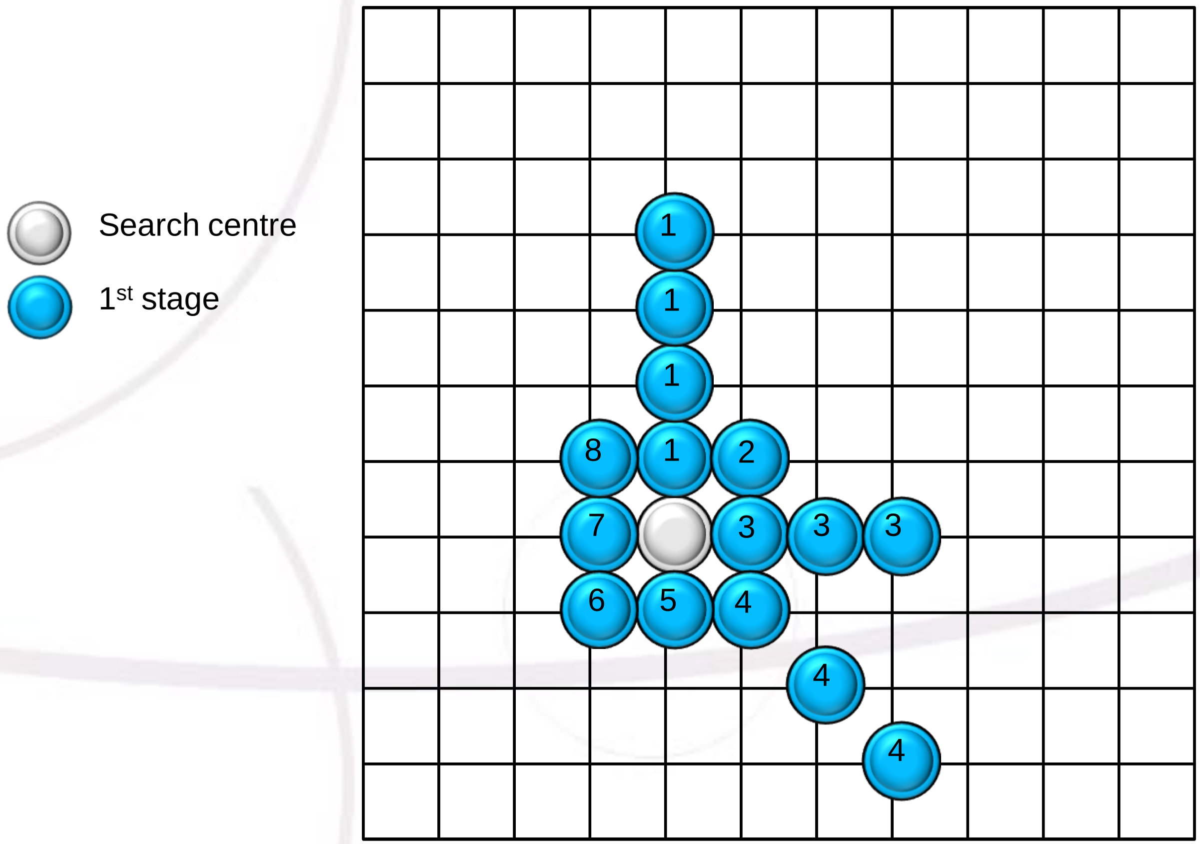
# Hexagonal Block Search (HEXBS)

- Each new stage (except the reduced step stage) evaluates three blocks
- It is faster than the DS, but has a lower quality of prediction




C.-H. Cheung and L.-M. Po, "Novel cross-diamond-hexagonal search algorithms for fast block motion estimation," IEEE Trans. Multimedia, vol. 7, no. 1, 2005

# Multi-Directional Gradient Descent Search (MDGDS)

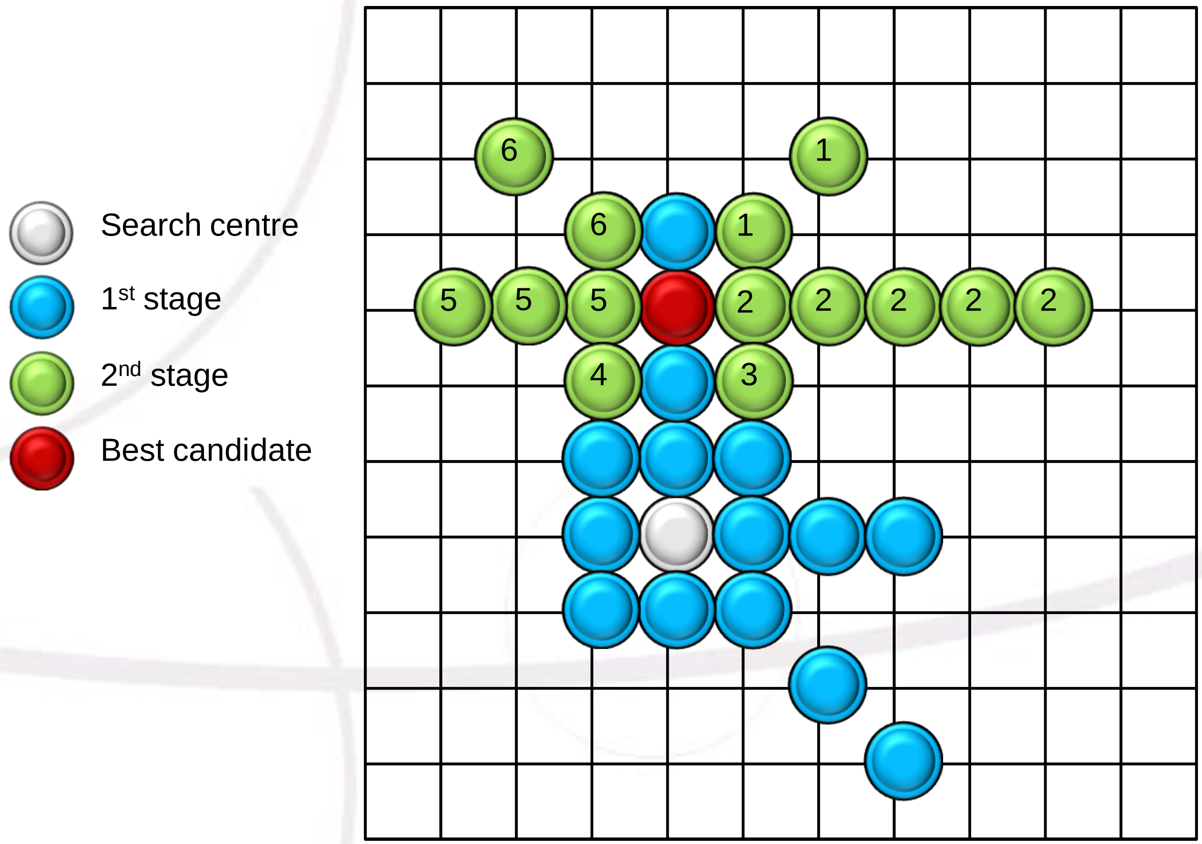



L.-M. Po, K.-H. Ng, K.-M. Wong, and K.-W. Cheung, “Multi-direction search algorithm for block-based motion estimation,” in IEEE Asia Pacific Conf. in Circuits and Systems (APPCAS), 2008

Universidad del Valle

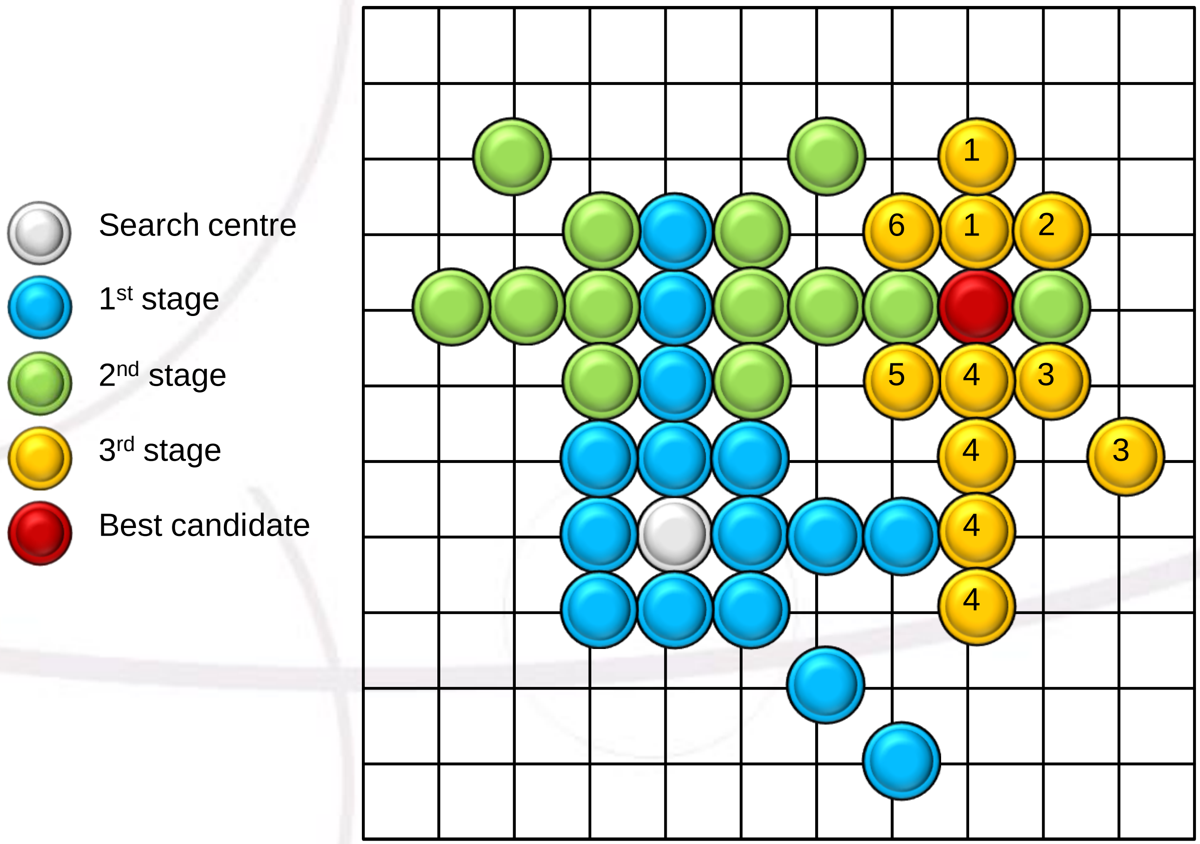
# Multi-Directional Gradient Descent Search (MDGDS)

Laboratorio de Multimedia y Vision




L.-M. Po, K.-H. Ng, K.-M. Wong, and K.-W. Cheung, “Multi-direction search algorithm for block-based motion estimation,” in IEEE Asia Pacific Conf. in Circuits and Systems (APPCAS), 2008

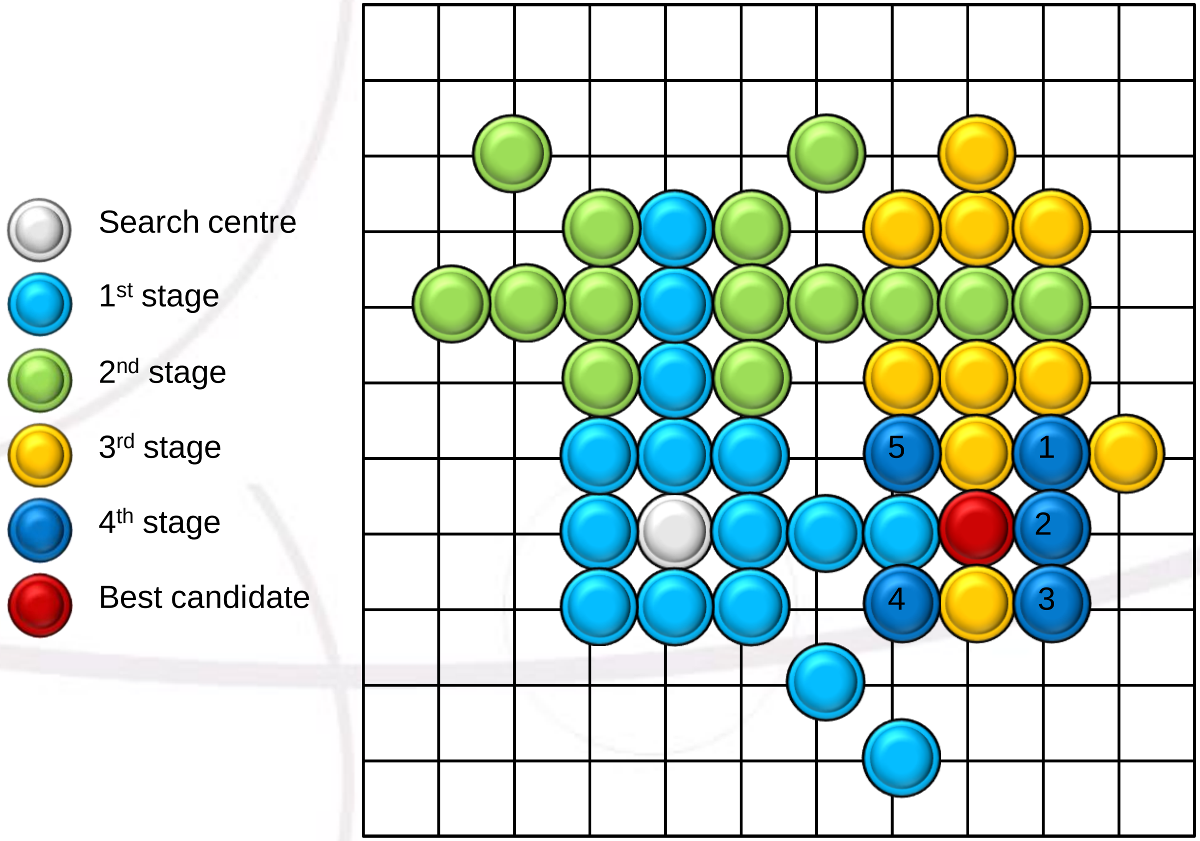
# Multi-Directional Gradient Descent Search (MDGDS)




L.-M. Po, K.-H. Ng, K.-M. Wong, and K.-W. Cheung, “Multi-direction search algorithm for block-based motion estimation,” in IEEE Asia Pacific Conf. in Circuits and Systems (APPCAS), 2008

# Multi-Directional Gradient Descent Search (MDGDS)

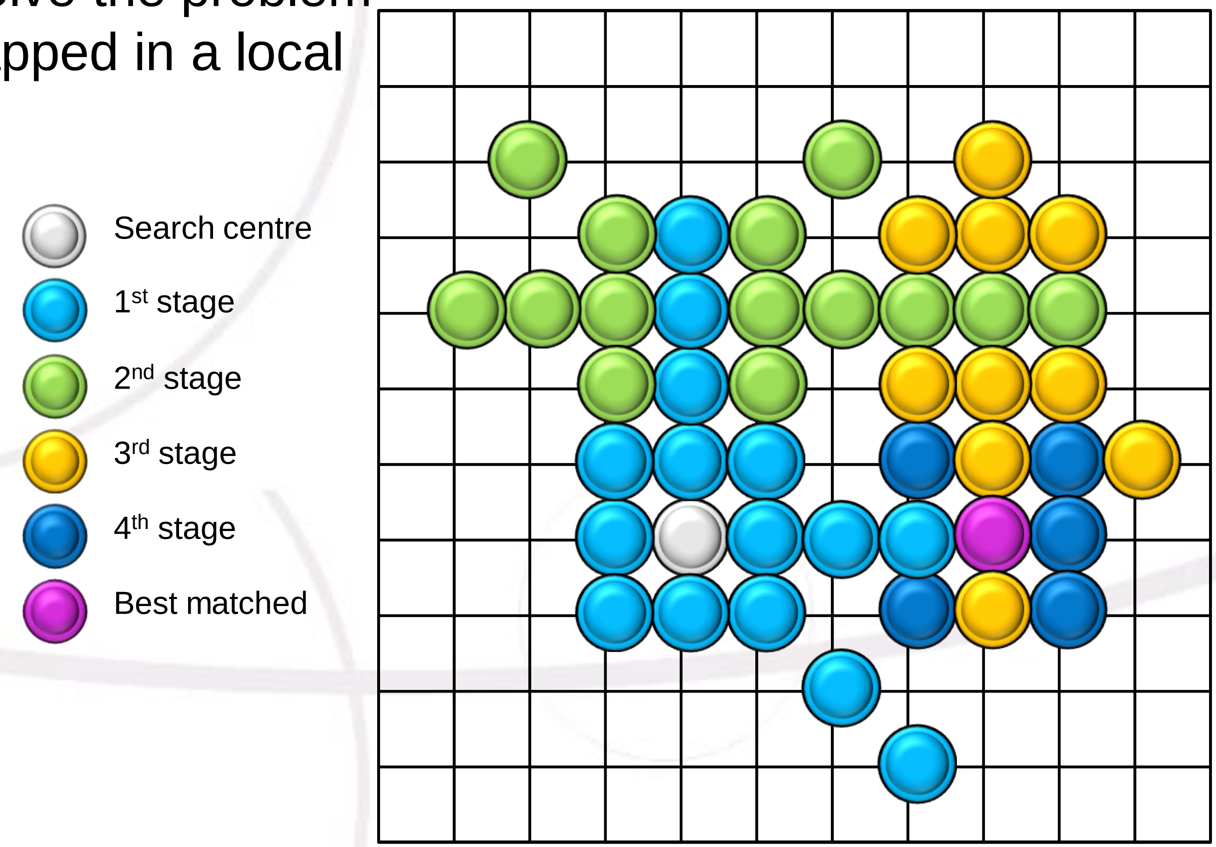





L.-M. Po, K.-H. Ng, K.-M. Wong, and K.-W. Cheung, "Multi-direction search algorithm for block-based motion estimation," in IEEE Asia Pacific Conf. in Circuits and Systems (APPCAS), 2008

# Multi-Directional Gradient Descent Search (MDGDS)

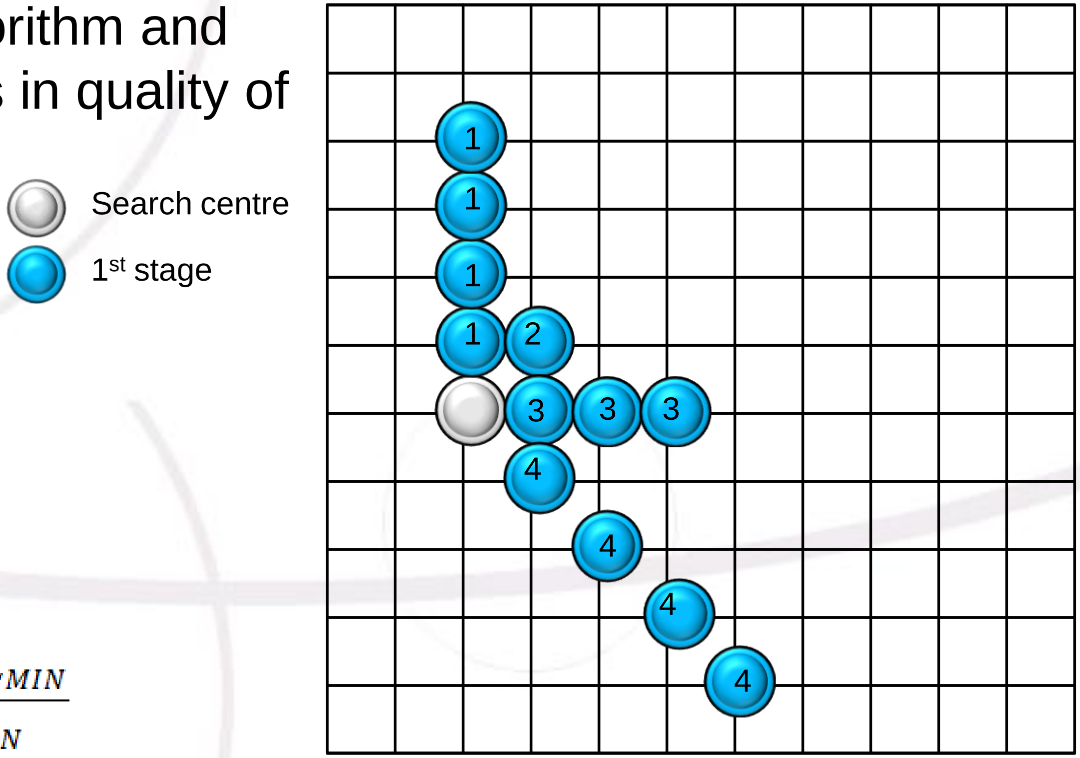
- It tries to solve the problem of being trapped in a local minimum




L.-M. Po, K.-H. Ng, K.-M. Wong, and K.-W. Cheung, “Multi-direction search algorithm for block-based motion estimation,” in IEEE Asia Pacific Conf. in Circuits and Systems (APPCAS), 2008

# Fast Directional Gradient Descent Search (FDGDS)

- It is an improvement of the MDGDS that increases the speed of the algorithm and leads to little loss in quality of prediction

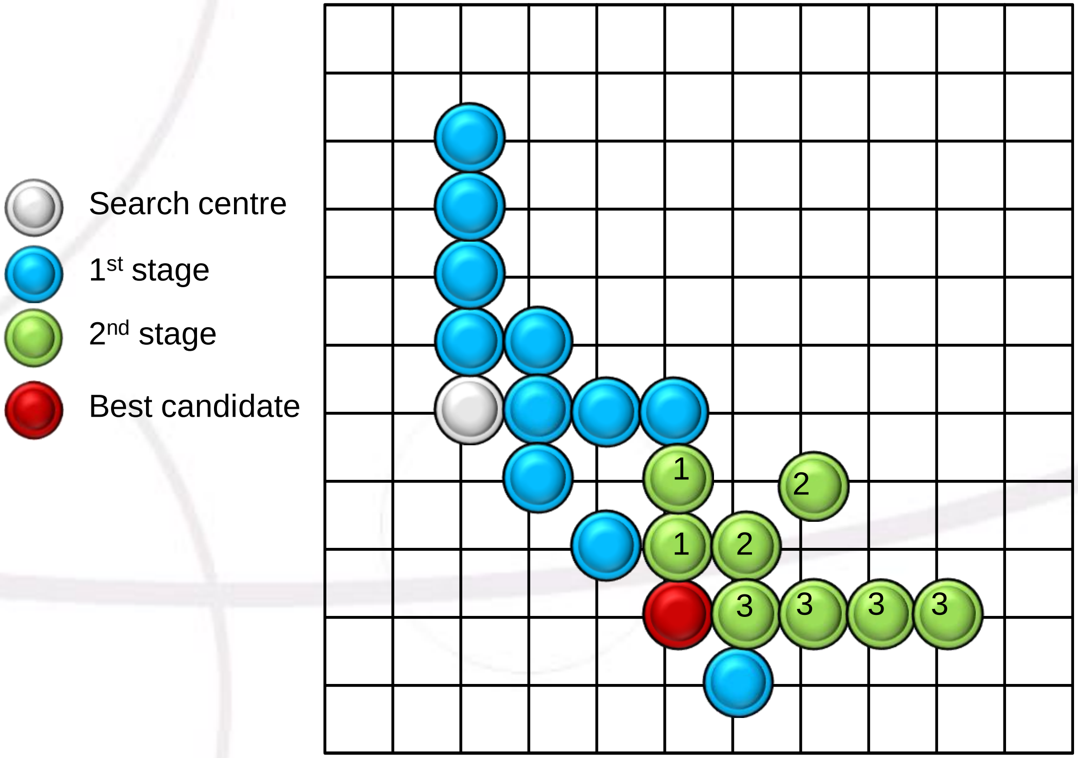



**Relative Distortion Ratio**

$$RDR = \frac{DIRECTIONAL_{MIN}}{CURRENT_{MIN}}$$

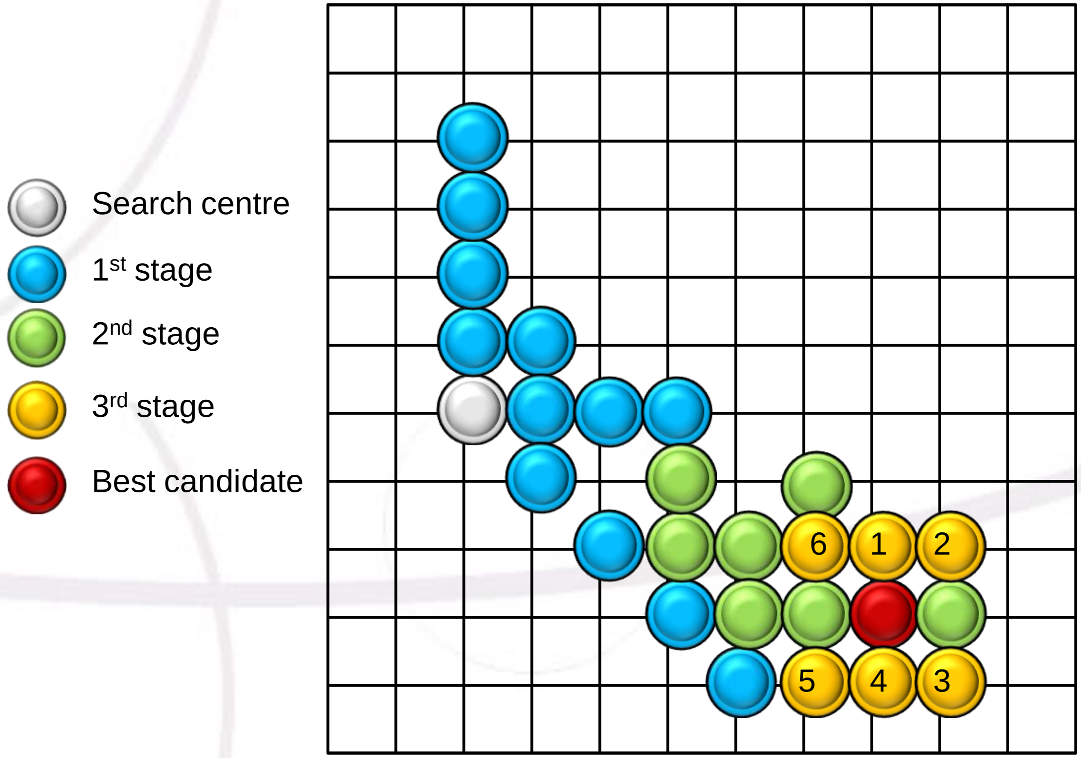
L.-M. Po, K.-H. Ng, K.-W. Cheung, K.-M. Wong, Y. Uddin, and C.-W. Ting, "Novel Directional Gradient Descent Searches for Fast Block Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 8, 2009

# Fast Directional Gradient Descent Search (FDGDS)



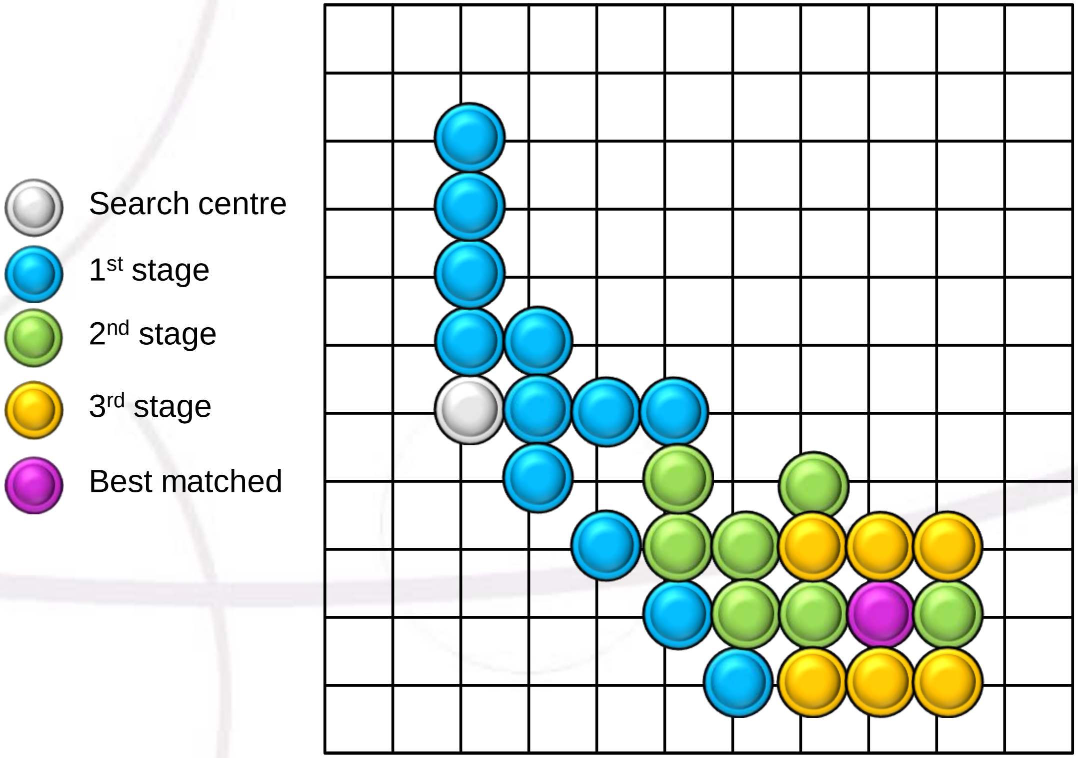

L.-M. Po, K.-H. Ng, K.-W. Cheung, K.-M. Wong, Y. Uddin, and C.-W. Ting, "Novel Directional Gradient Descent Searches for Fast Block Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 8, 2009

# Fast Directional Gradient Descent Search (FDGDS)




L.-M. Po, K.-H. Ng, K.-W. Cheung, K.-M. Wong, Y. Uddin, and C.-W. Ting, “Novel Directional Gradient Descent Searches for Fast Block Motion Estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 8, 2009

# Fast Directional Gradient Descent Search (FDGDS)




L.-M. Po, K.-H. Ng, K.-W. Cheung, K.-M. Wong, Y. Uddin, and C.-W. Ting, "Novel Directional Gradient Descent Searches for Fast Block Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 8, 2009

# Quality Metrics

## Peak Signal-to-Noise Ratio

$$MSE(X,Y) = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[X(i,j) - Y(i,j)]^2$$

$$PSNR(X,Y) = 20\log_{10}\left(\frac{MAX_Y}{\sqrt{MSE(X,Y)}}\right)$$

- It is a point to point metric
- Based on square differences
- It is not very well matched to perceived visual quality

## Structural Similarity Index

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

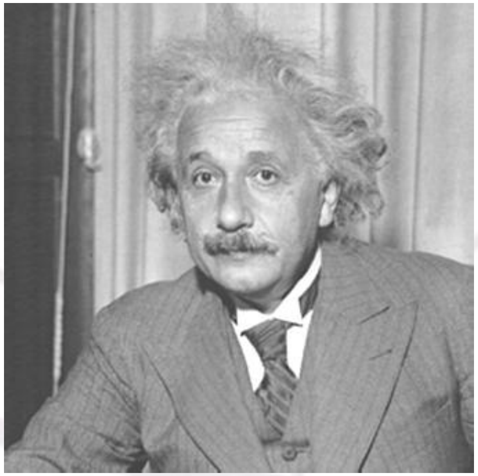
$$MSSIM(X,Y) = \frac{1}{n}\sum_{i=0}^{n-1} SSIM(x_i, y_i)$$

- It is a windowed metric
- Based on luminance, contrast and structure between an original and a distorted images
- It takes into account the visual perception of the image

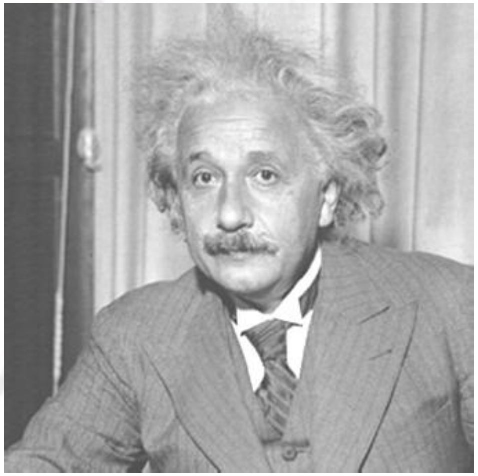
Z. Wang, A. C. Bovik, H. R. Sheikh and, E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, 2004

C. S. varnan, A. Jagan, J. Kaur, D. Jyoti, and D. S. Rao, "Image quality assessment techniques pn spatial domain," International Journal on Computer Science and Technology, vol. 2, no. 3, 2011

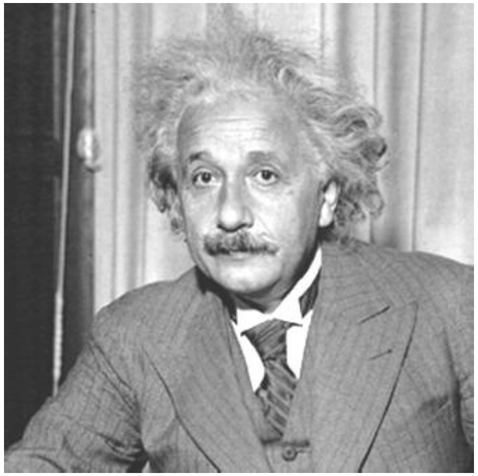
# Quality Metrics (ii)



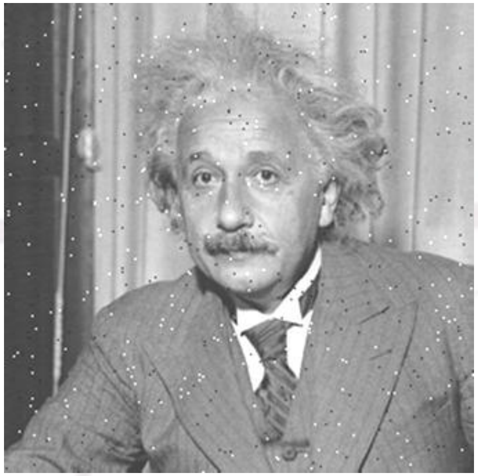
Original image
PSNR undefinied, SSIM = 1



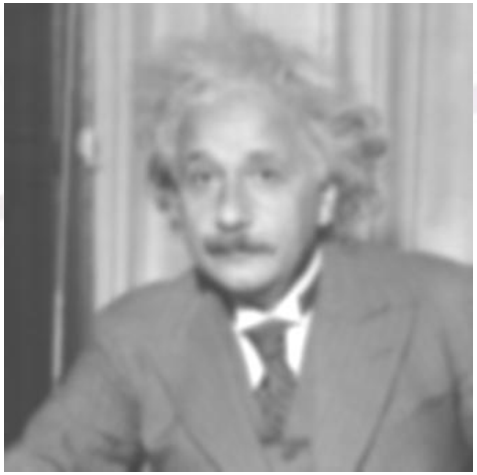
PSNR = 26.547, SSIM = 0.988



PSNR = 26.547,  SSIM = 0.913



PSNR = 26.547,  SSIM = 0.840



PSNR = 26.547,  SSIM = 0.694

https://ece.uwaterloo.ca/~z70wang/research/ssim/

# Performance Metrics

- Since an algorithm requires time proportional to the number of explored blocks (EXB), the computational cost of a BMA is determined by the EXB

**EXB in the case of Zero Motion Vector (ZMV)**

| BMA | 3SS | 4SS | DS | HEXBS | MDGDS | FDGDS |
|---|---|---|---|---|---|---|
| EXB | 25 | 17 | 13 | 11 | 9 | 9 |

V. Padilla, “Algoritmos de block-matching para compresión de video,” Final Career Project, Systems Engineering Program, Universidad del Valle, 2009

Universidad del Valle

Laboratorio de Multimedia y Vision

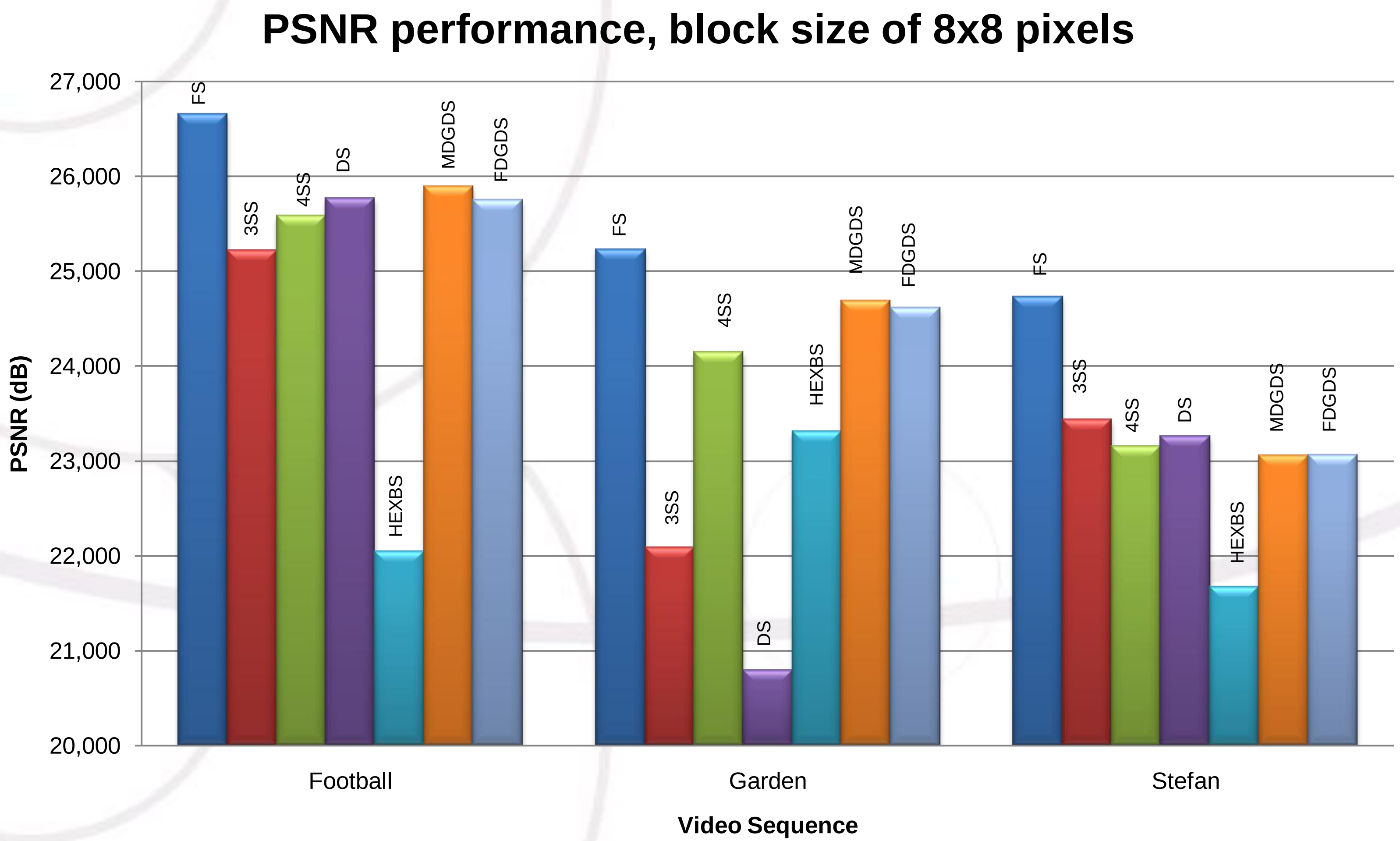
# Video Test Sequences

| Sequence | Size | # Frames | Motion |
|---|---|---|---|
| Akiyo | 352x288 | 300 | Small |
| Mother_ daugthter | 352x288 | 300 | Small |
| Silent | 352x288 | 300 | Small |
| Foreman | 352x288 | 300 | Medium |
| Garden | 352x240 | 115 | Medium |
| Mobile | 352x288 | 300 | Medium |
| Coastguard | 352x288 | 300 | Large |
| Football | 352x288 | 260 | Large |
| Stefan | 352x240 | 300 | Large |

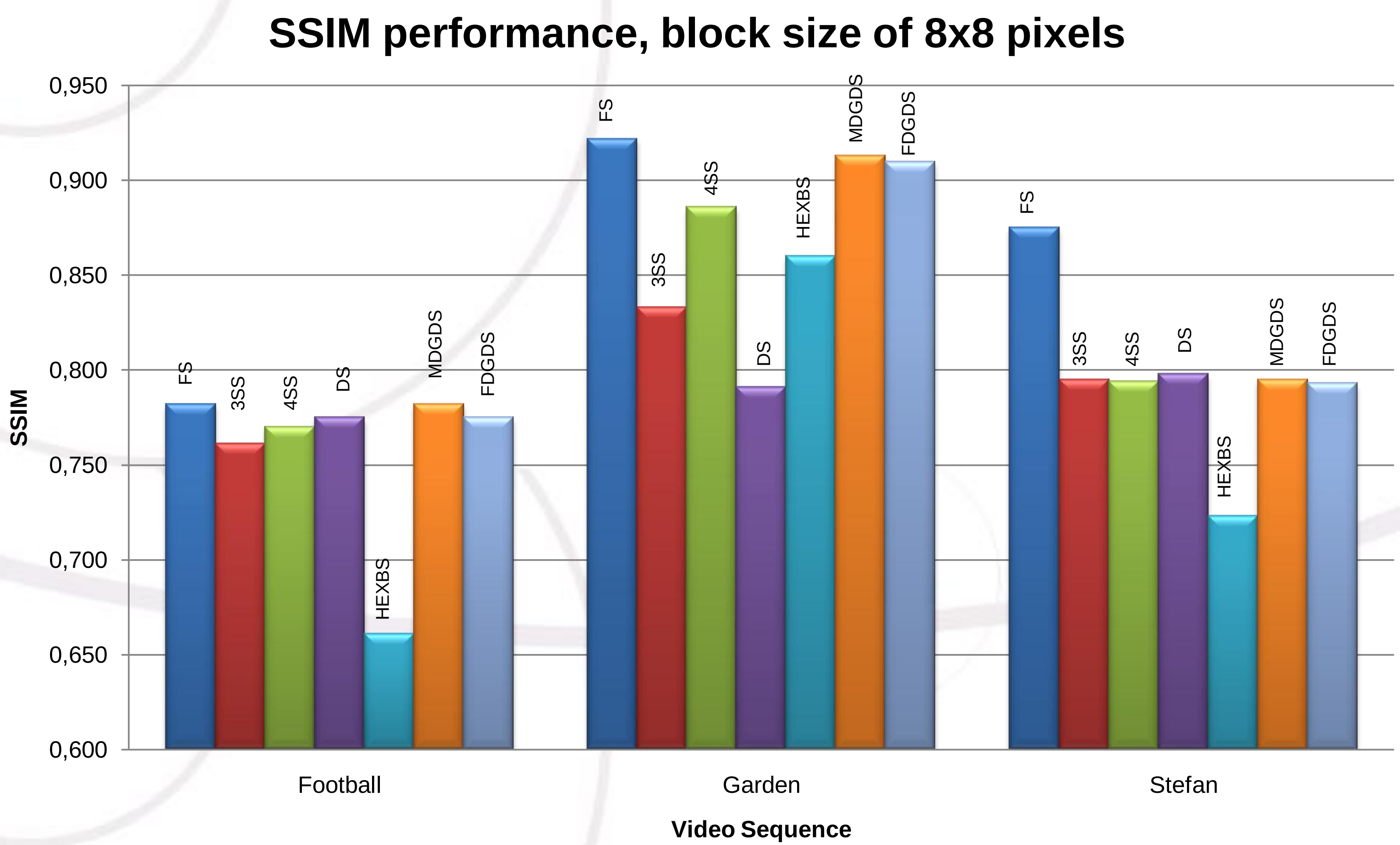

Block sizes used: 8x8, 16x16 and 32x32

All video sequences used are in uncompressed format: YUV4MPEG, and are available at: http://media.xiph.org/video/derf/

# Results

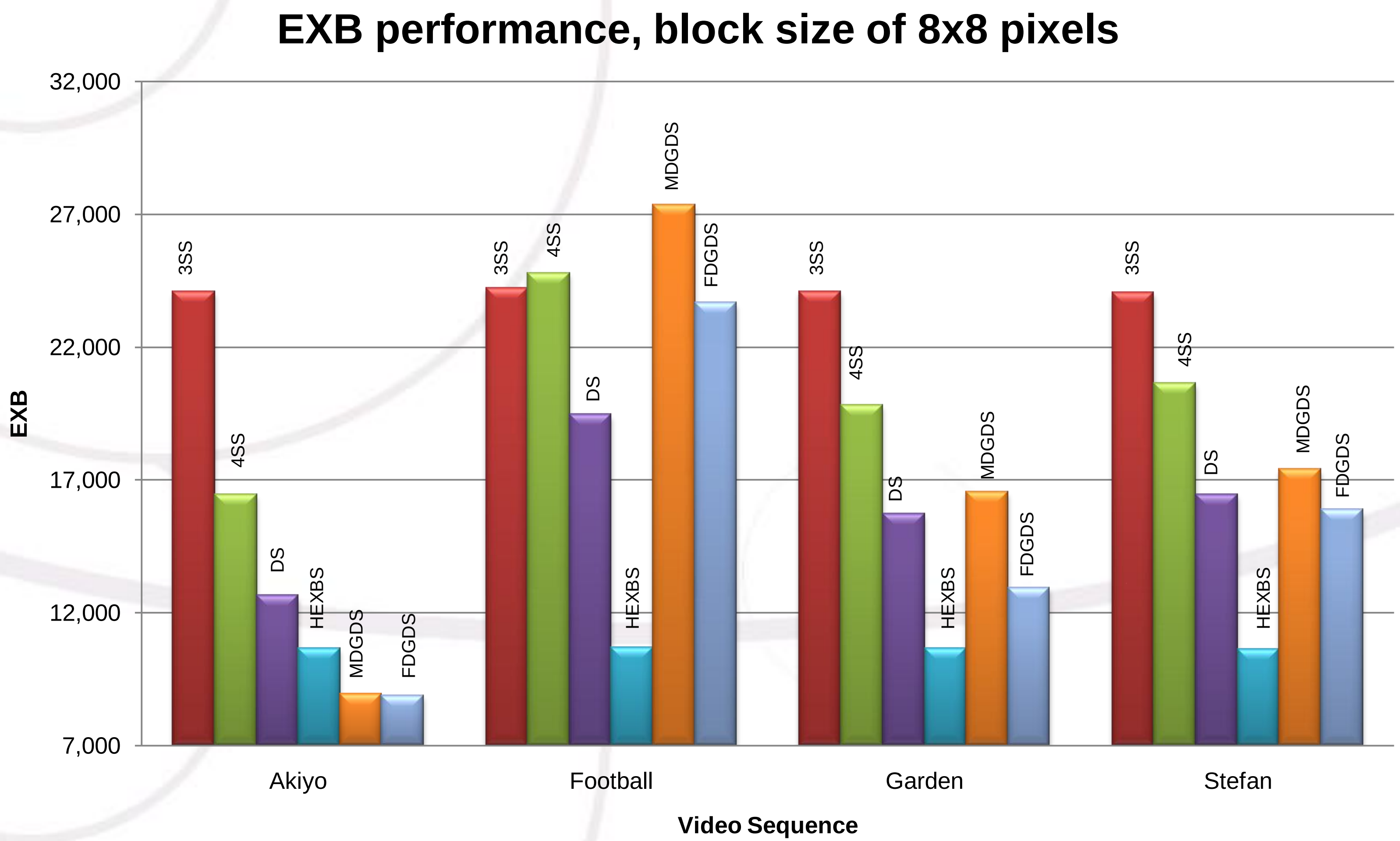

PSNR performance, block size of 8x8 pixels
27,000
26,000
25,000
24,000
23,000
22,000
21,000
20,000
PSNR (dB)
FS
3SS
4SS
DS
HEXBS
MDGDS
FDGDS
Football
Garden
Stefan
Video Sequence

# Results (ii)

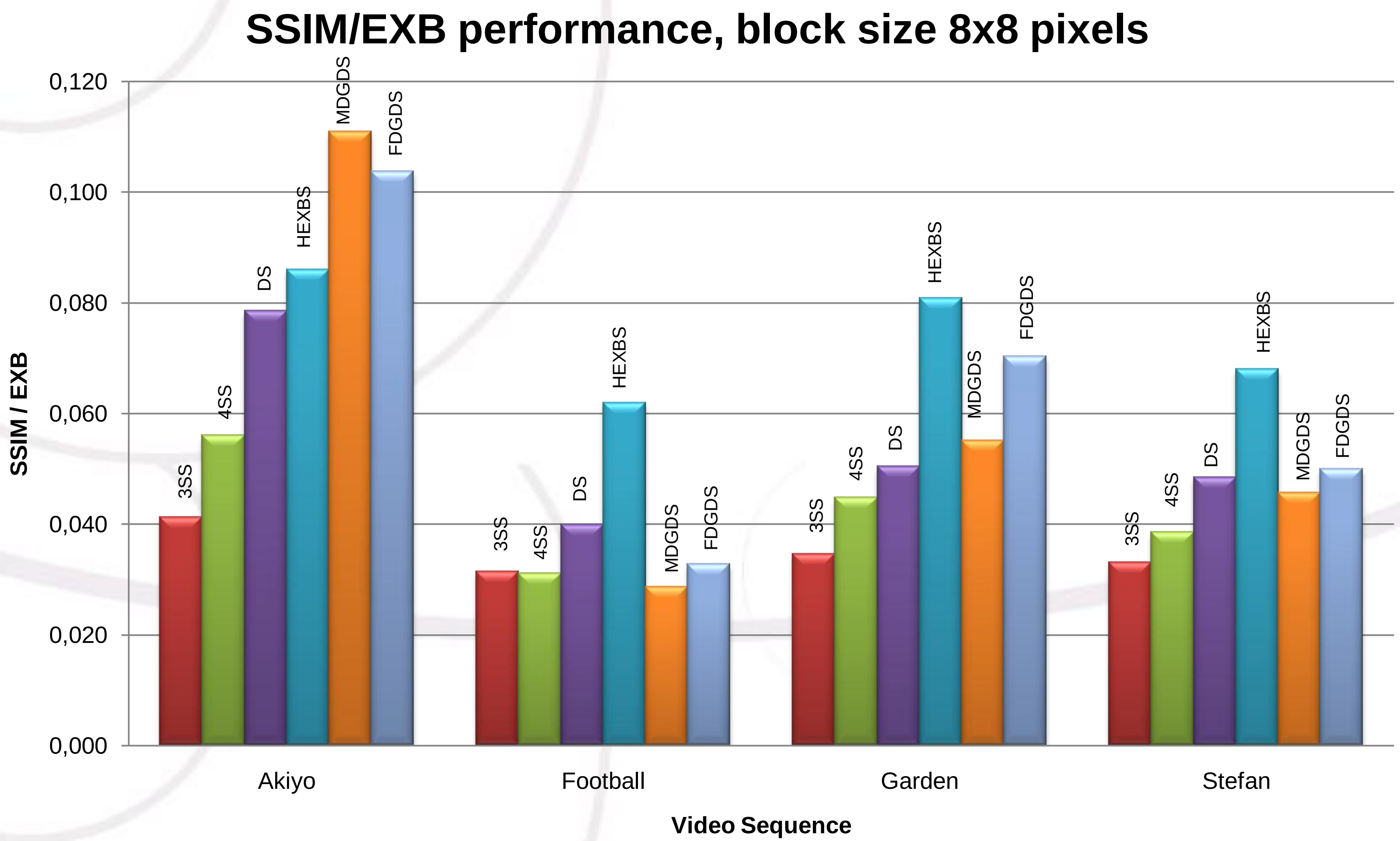


SSIM performance, block size of 8x8 pixels
0,950
0,900
0,850
0,800
0,750
0,700
0,650
0,600
SSIM
FS
3SS
4SS
DS
HEXBS
MDGDS
FDGDS
Football
Garden
Stefan
Video Sequence

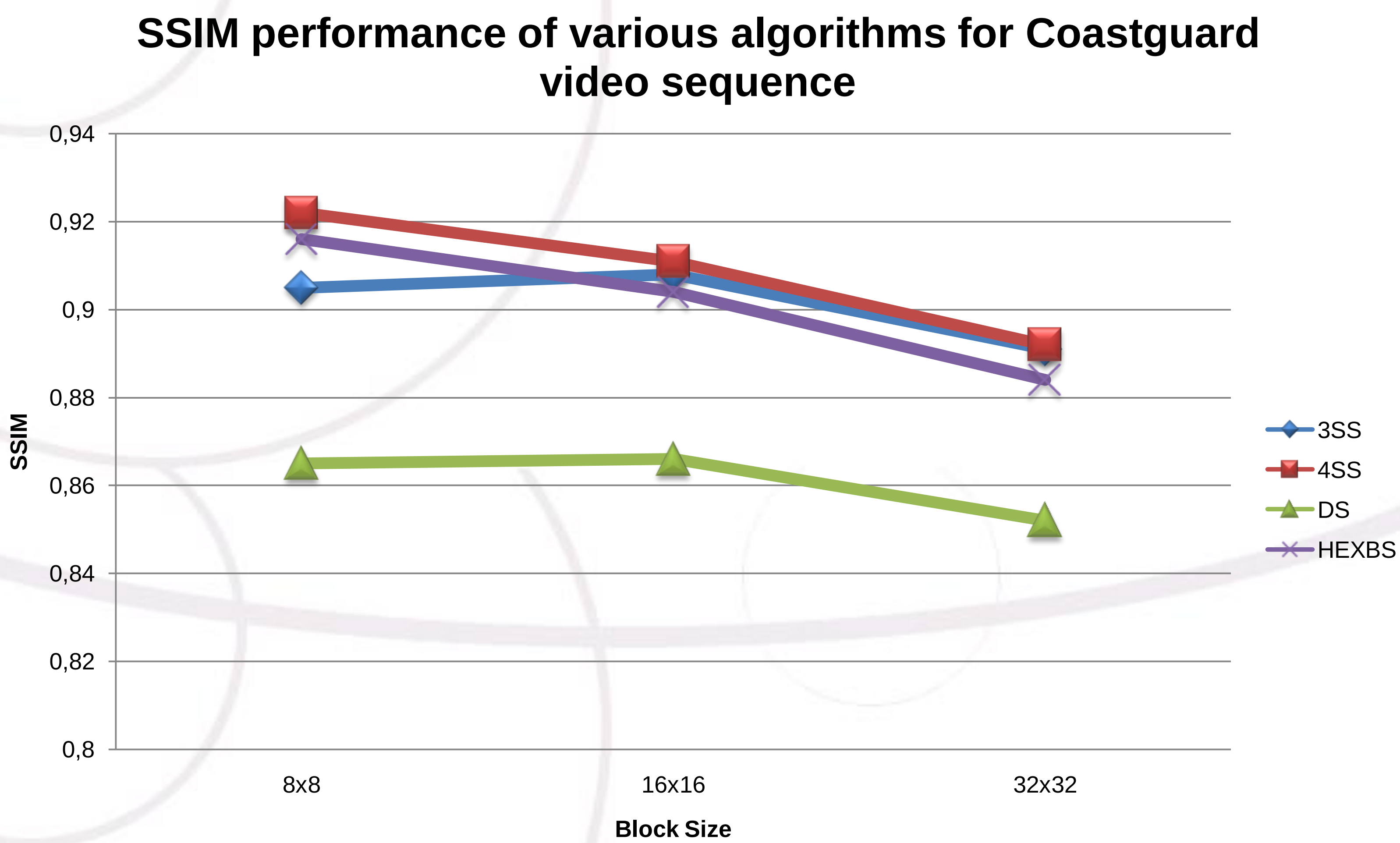
# Results (iii)


EXB performance, block size of 8x8 pixels
32,000
27,000
22,000
17,000
12,000
7,000
EXB
3SS
4SS
DS
HEXBS
MDGDS
FDGDS
Akiyo
Football
Garden
Stefan
Video Sequence

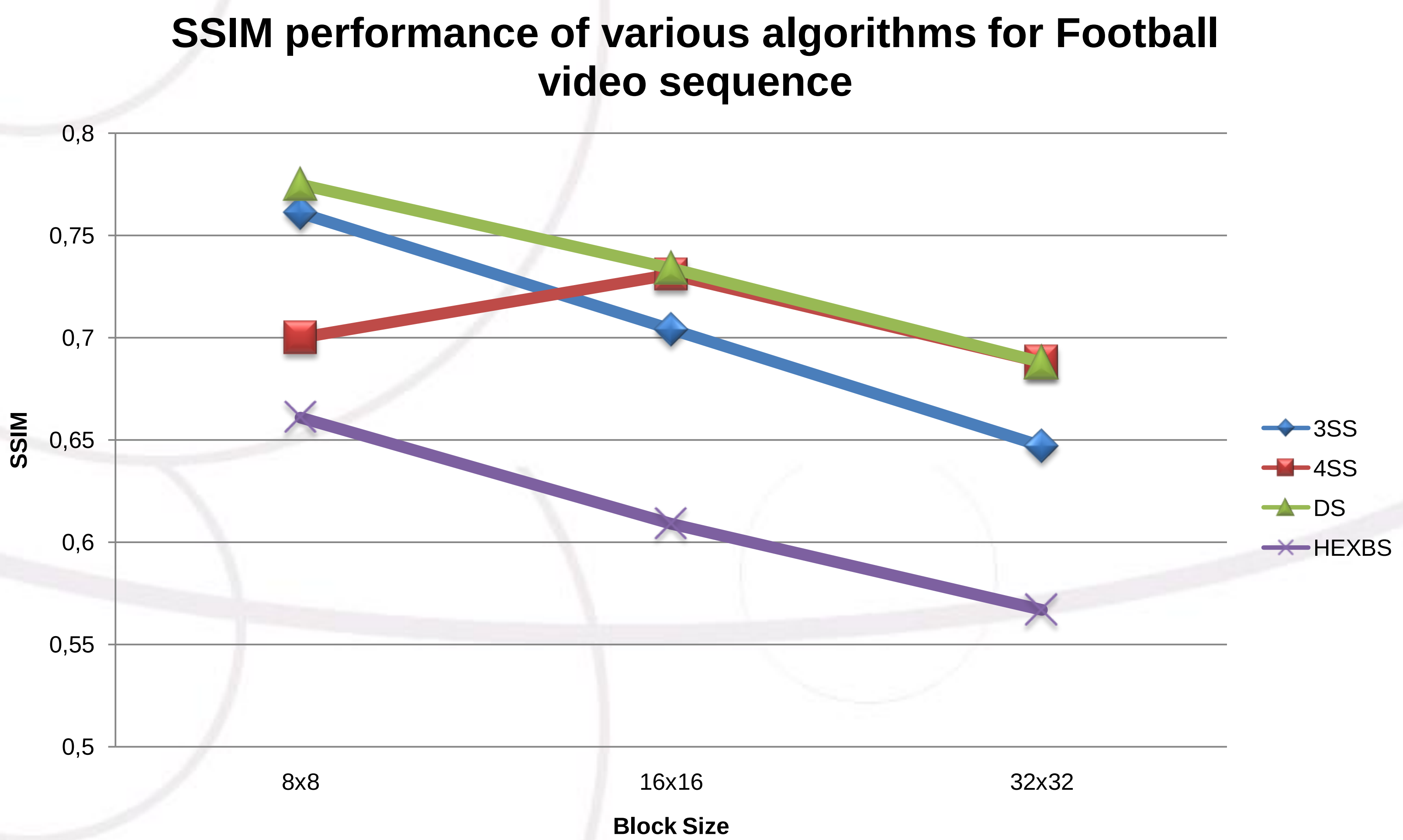
# Results (iv)


SSIM/EXB performance, block size 8x8 pixels
SSIM / EXB
0,120
0,100
0,080
0,060
0,040
0,020
0,000
3SS
4SS
DS
HEXBS
MDGDS
FDGDS
Akiyo
Football
Garden
Stefan
Video Sequence

# Results (v)

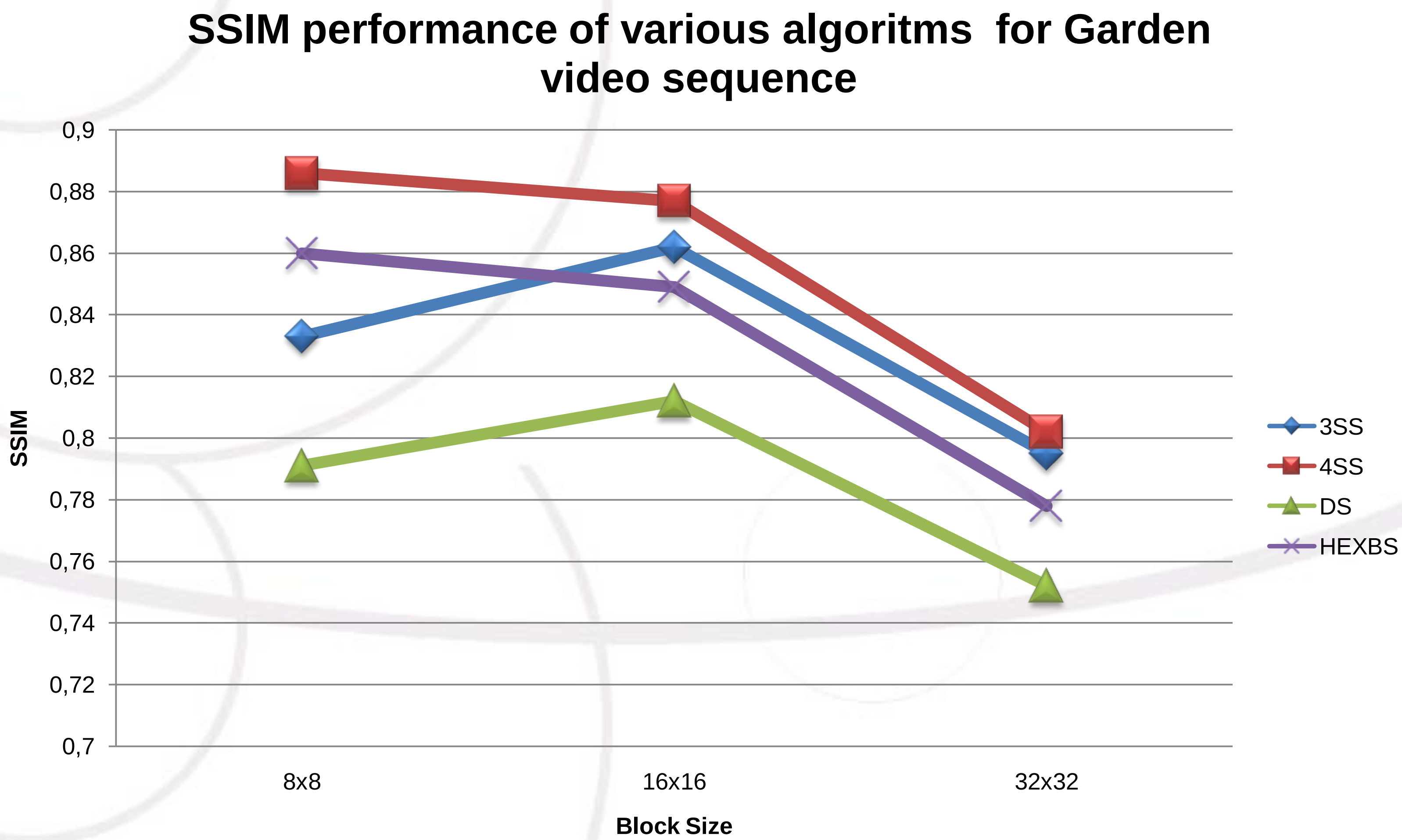

SSIM performance of various algorithms for Coastguard video sequence
0,94
0,92
0,9
0,88
0,86
0,84
0,82
0,8
SSIM
8x8
16x16
32x32
Block Size
3SS
4SS
DS
HEXBS

# Results (vi)



SSIM performance of various algorithms for Football video sequence
0,8
0,75
0,7
0,65
0,6
0,55
0,5
SSIM
8x8
16x16
32x32
Block Size
3SS
4SS
DS
HEXBS

# Results (vii)



SSIM performance of various algoritms for Garden video sequence
0,9
0,88
0,86
0,84
0,82
0,8
0,78
0,76
0,74
0,72
0,7
SSIM
8x8
16x16
32x32
Block Size
3SS
4SS
DS
HEXBS

Universidad del Valle

# Final Remarks

Laboratorio de Multimedia y Vision

- The HEXBS shows low computational cost but produces low quality of prediction
- The MDGDS and the FDGDS show low computational cost and produce the highest quality of prediction
- The FGDGS achieves a good trade off between high quality of prediction and a low computation cost
- The HEXBS is less affected by the variation in the block sizes, whilst the others show a big loss of prediction by increasing the block size used

# THANKS!