Complexity Reduction for Video Surveillance Systems

Hyuk-Jae Lee

Seoul National University

2016. 12. 05

Contents

- Intelligent Video Cameras
- Power consumption in video cameras
- Reduction of complexity/power consumption
- Future trends

Video Cameras

CC-TVs

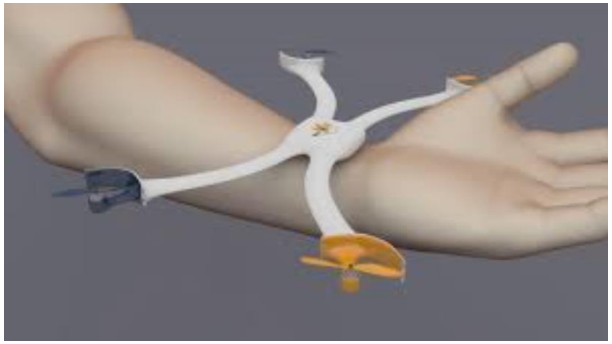
Inside car

Drone camera

sports camera

Badge camera

Video Cameras (2): Selfie Drone



Video Cameras (3)

- Life logger: continuously record visual experience
- 1 Mega bit/second x ~10 Mega seconds/year = ~ 10 Tera bits/year

Video Analysis in CC-TV

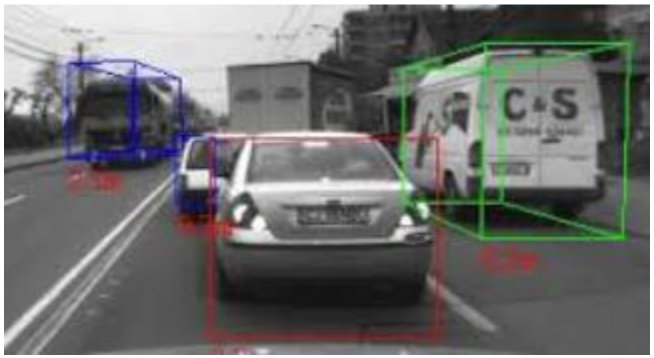
Fall detection

Intrusion/thief detection

violence detection

Video Analysis by camera for a car

car detection



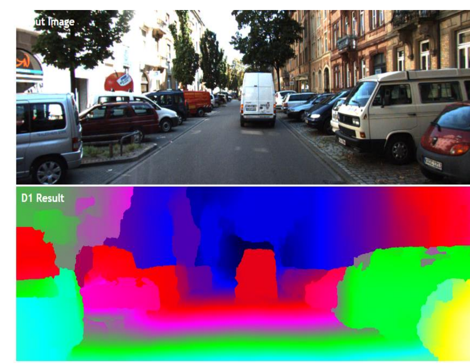
pedestrian detection

drowsiness

Traffic signal detection

lane detection

distance detection



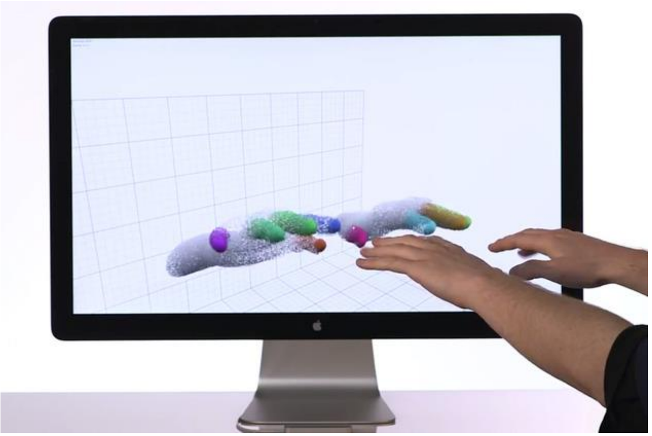
Human detection application

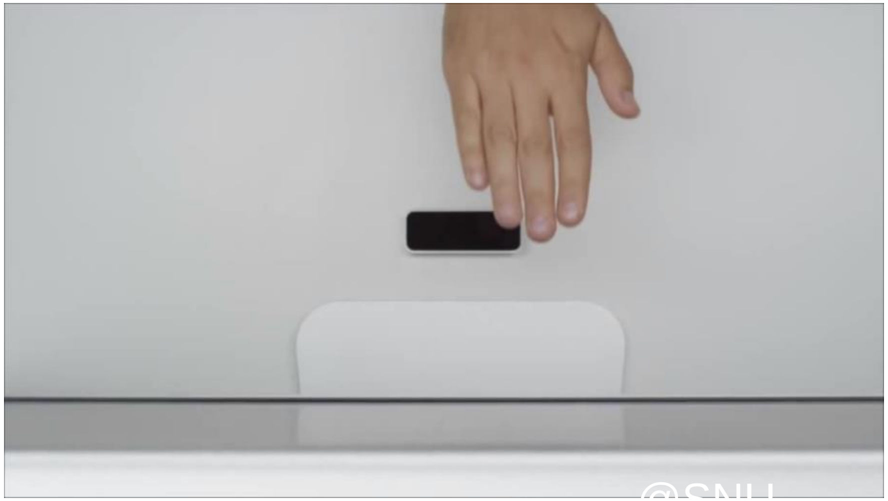
Head count analysis

Digital signage:
estimate its effectiveness

Video mining: in-store behavior analysis

Hand gesture recognition





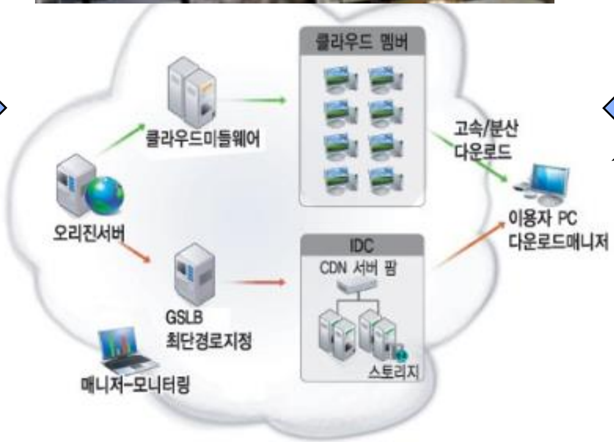
Contents

- Intelligent Video Cameras
- Power consumption in video cameras
- Reduction of complexity/power consumption
- Future trends

CC-TV Data

- What data to store? What data to process?
- Where to store ? Where to process?

Cloud center ?

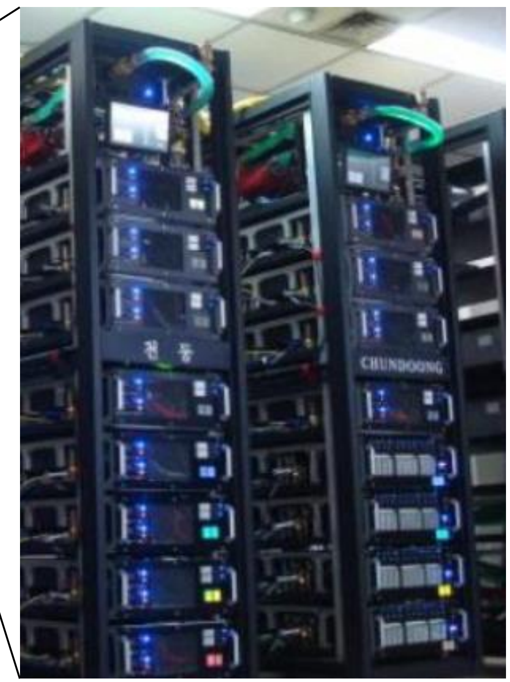


Power consumption of Big Data Processing

Go competition: Human vs AI (2016. 3)

20 W

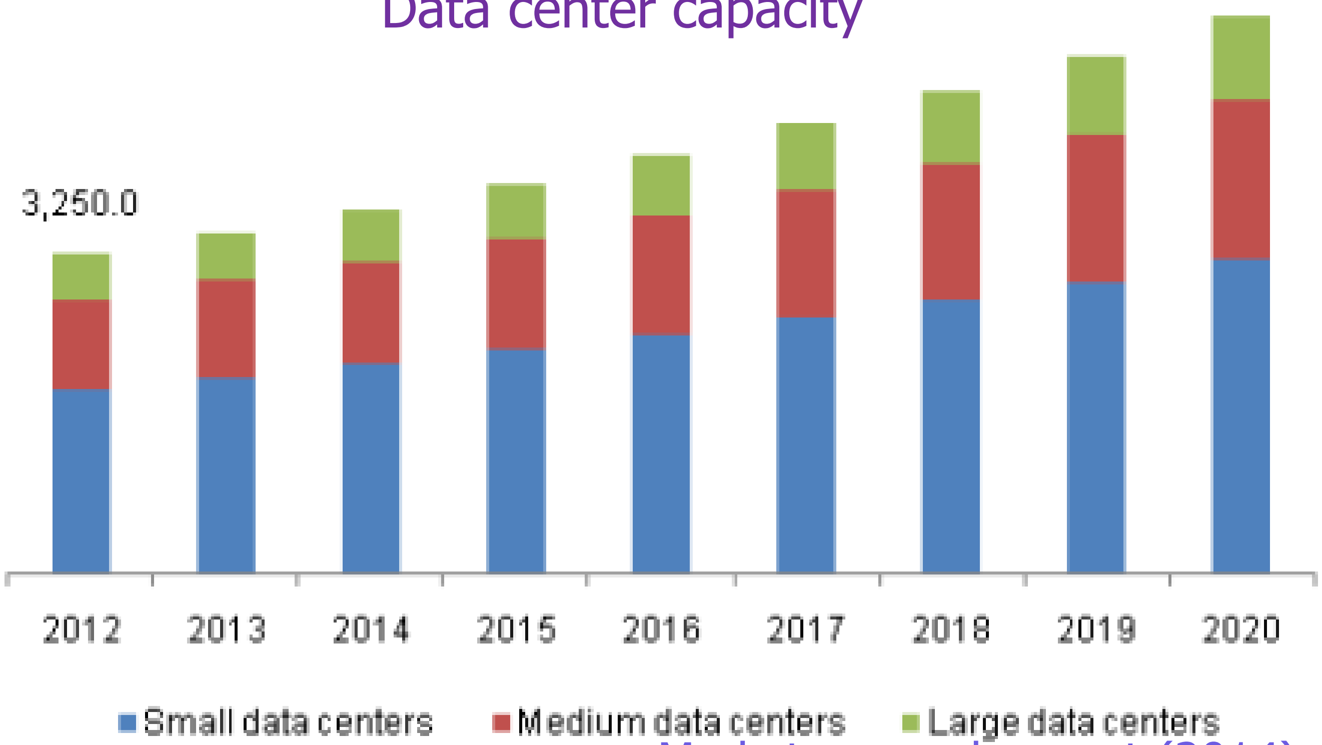
170,000 W (CPU: 1202, GPU 176)



Power Consumption by Data Centers

- 70 billion kilo watts-hour consumed by data centers in the US (2014)
  - 2% of total power consumption of US
  - Equivalent to 6.4 million homes

Data center capacity

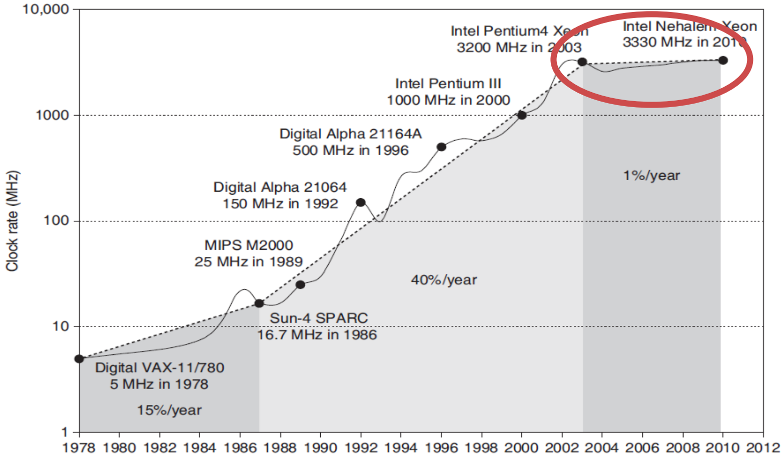


Market research report (2014)

Technology Evolution

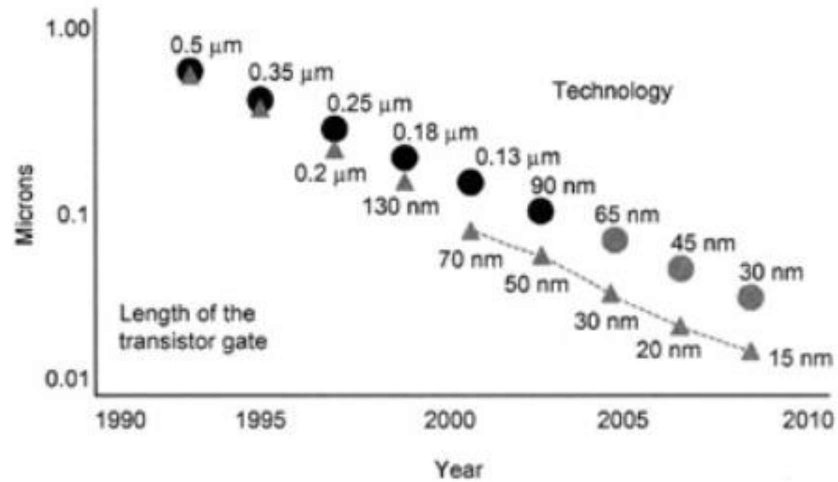
CPU Speed

Limited by
Power consumption

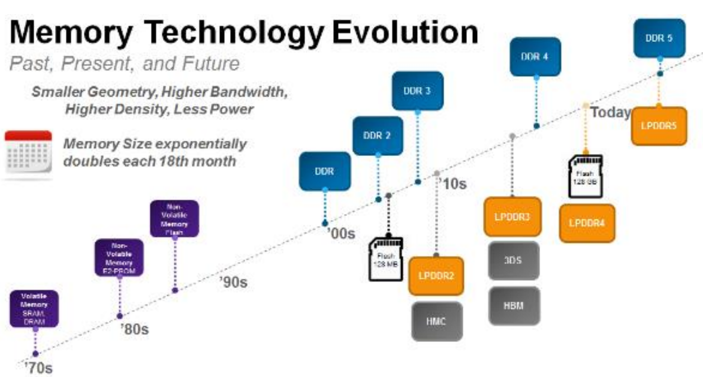


2003

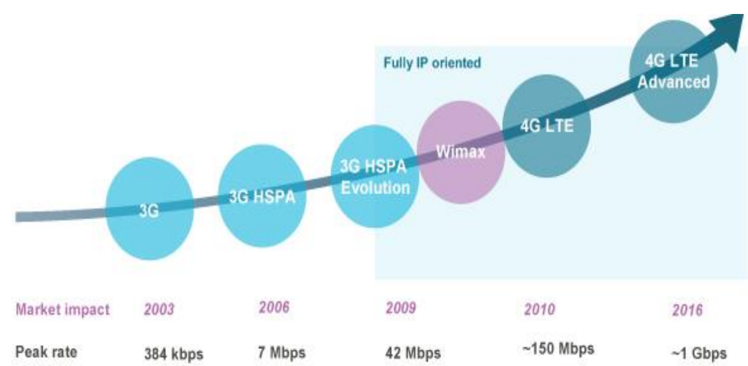
Semiconductor Technology Shrink



DRAM capacity



Communication bandwidth



Contents

- Intelligent Video Cameras
- Power consumption in video cameras
- Reduction of complexity/power consumption
- Future trends

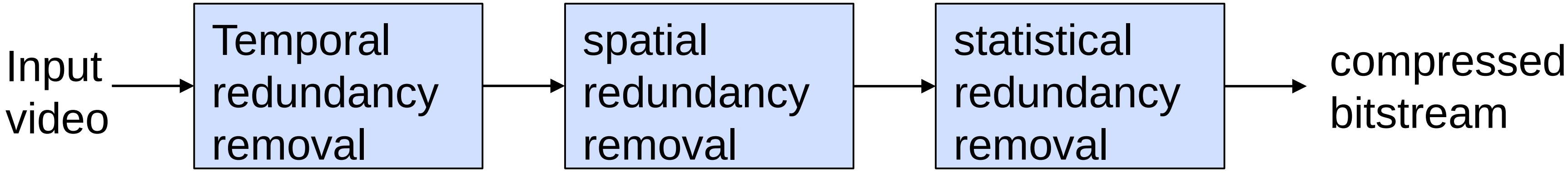
CC-TV Data

- What data to store? What data to process?
- Where to store ? Where to process?

Cloud center ?

Video compression

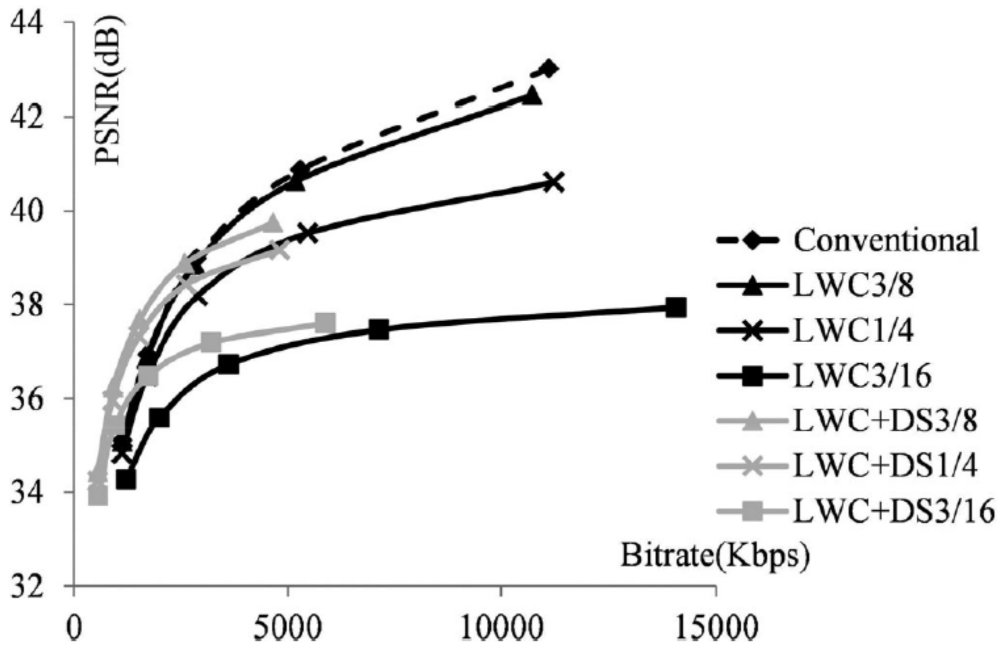
- Reduce the amount of data by 1/50 ~ 1/200
  - Reduce the storage space for local storage or
  - reduce communication bandwidth for remote data transfer
- Basic idea: remove redundant data → reconstruct the original data without the removed data



- Temporal redundancy removal
  - Compare with previous frames
  - Most effective
- Spatial redundancy removal
  - Compare with neighboring blocks in the same frame
- Statistical redundancy removal: use entropy coding

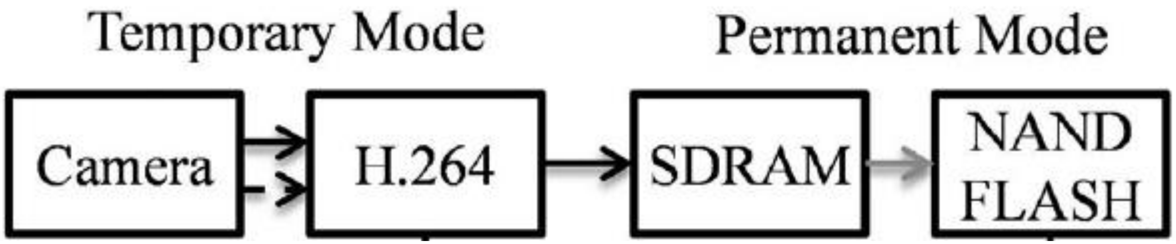
Evaluation of coding efficiency

- R-D (Rate-Distortion) Curve
  - PSNR (vertical axis): quality of video (estimation of distortion)
  - Bit rate (horizontal axis): data size (compression ratio)
- Trade-off between PSNR (distortion) and bit rate (data size)
- The coding efficiency is better if it is higher in all bit rate



Additional data reduction

- Temporary mode: camera → video compression → DRAM
  - Camera input captured and compressed continuously
- Permanent mode: DRAM → Flash memory
  - Important event detected, the DRAM data are stored in flash

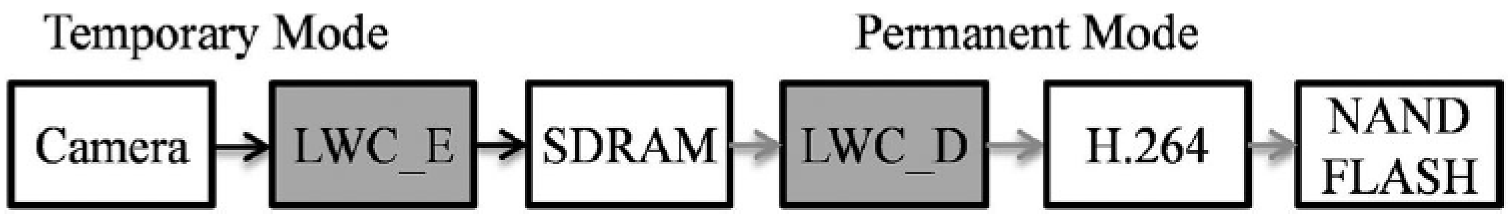


H.264: popular video compression standard

- Inefficiency:
  - video compression is complex demanding large power consumption

Computation complexity reduction

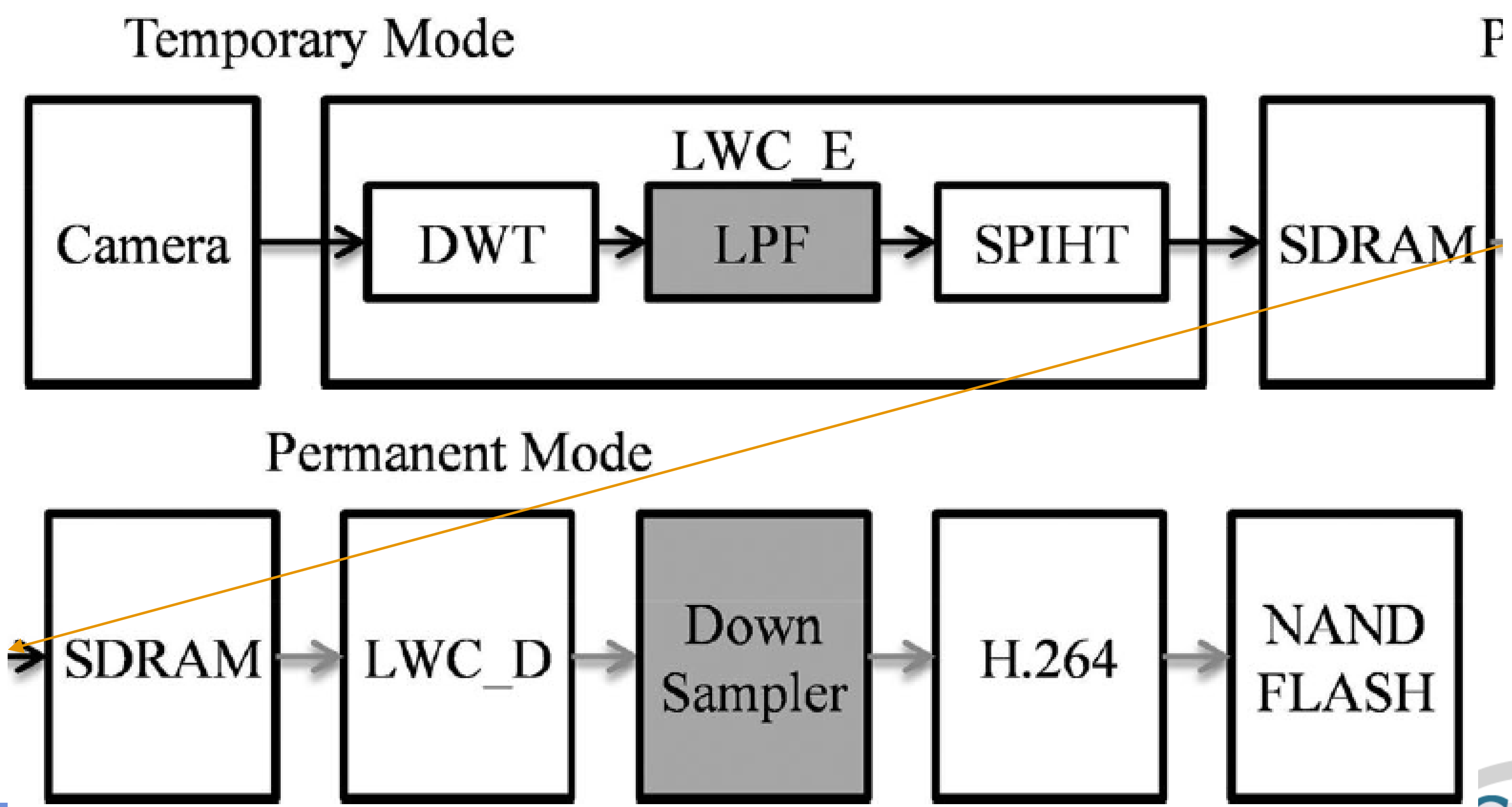
- Light-weight compression:
  - Less complex than standard video compression
  - Less efficient in compression
    - Target compression ratio: ½ - 1/10
    - Standard video compression: >> 1/10



- Temporary mode: Camera → LWC encoder → SDRAM
- Permanent mode: DRAM → LWC decoder → standard encoder → flash

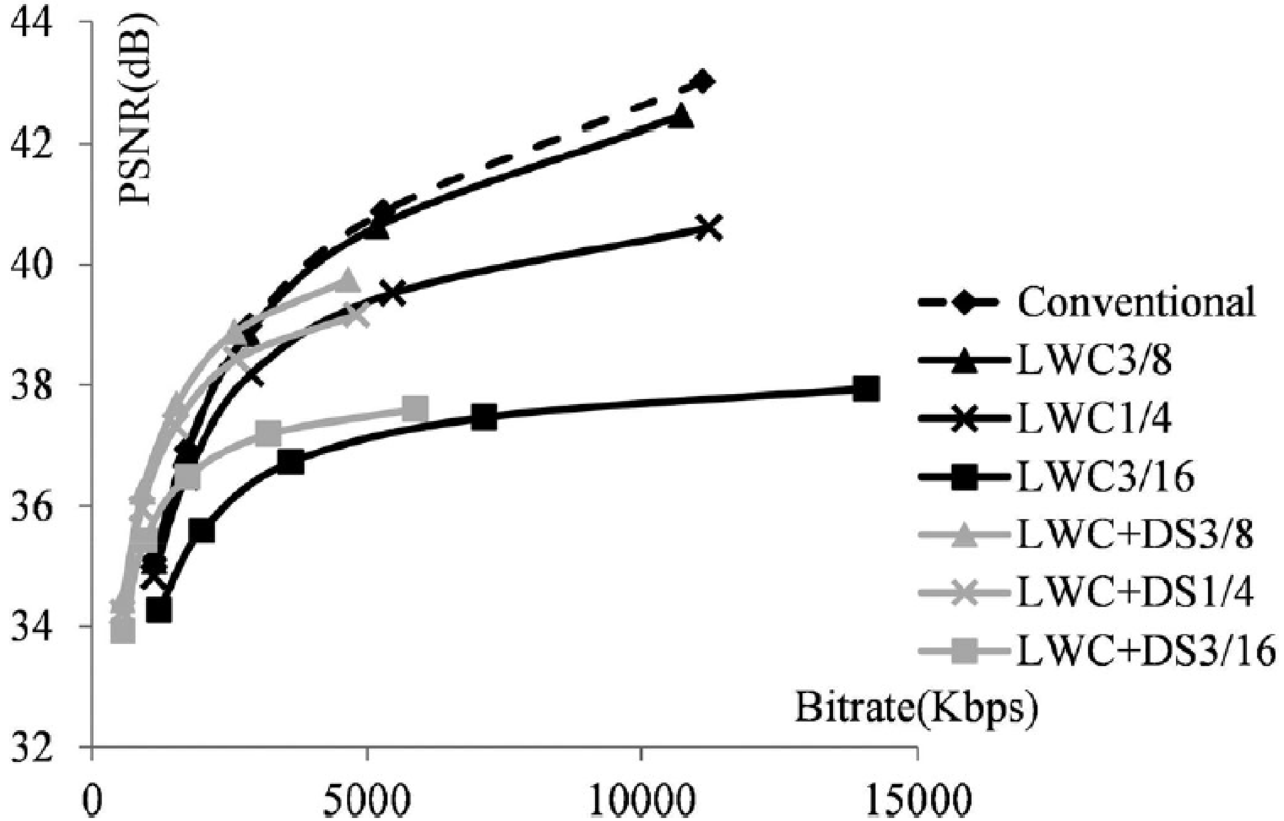
Reduction in the permanent mode

- LWC behaves as low-pass filter
- Down-sampling does not cause a significant loss of video quality
  - Reduced image size → reduced power consumption



Evaluation of down-sampling

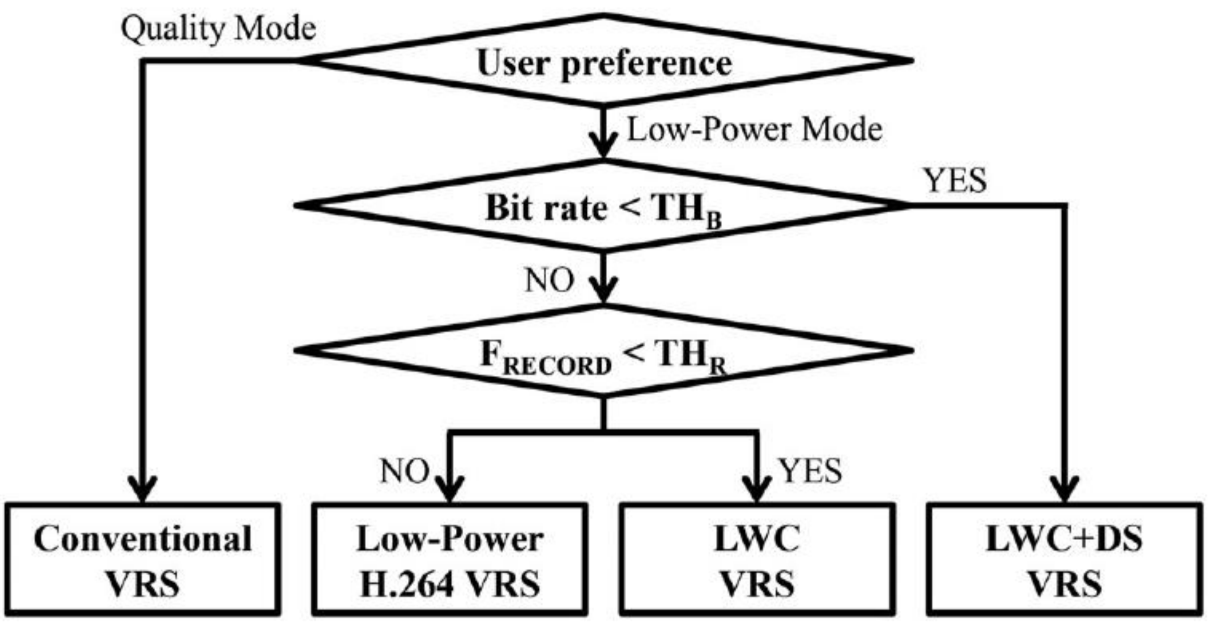
- Down-sampling some cases give higher quality
  - Effective for low bit rate



CAPP

Operating Mode Selection

- TH: Thresholds for mode decision
- Frecord: The frequency of the permanent mode
- Low power H.264: low-complexity H.264 with reduced compression efficiency
- LWC VRS: use of LWC in the temporary mode
- LPF+DS VRS: used of down-sampling in the permanent mode



Simulation Results

- Power consumption breakdown

POWER CONSUMPTION BY HARDWARE MODULE

	P_{264}	P_{LWCE}	P_{LWCD}	P_{DOWN}
Power (mW)	97.88	10.89	15.36	13.64

- External memory access

EXTERNAL BANDWIDTH FOR EACH COMPRESSION TYPE

	H.264/AVC	LWC	LWC + DS
Compression ratio	1/100	3/8	1/4
One pixel size (bits)	12	16	16
Out stream size per second (Mbits)	3.24	162	108
Memory write per second (Mbits)	648	0	0
Memory read per second (Mbits)	972	0	0

Simulation Results (2)

- Power consumption

POWER CONSUMPTION OF EACH MODE FOR EACH VRS

Video Recording Systems	Mode	P_{264} (mW)	P_{LWCE} (mW)	P_{LWCD} (mW)	P_{DOWN} (mW)	P_{SDRAM} (mW)	P_{NAND} (mW)	P_{MODE} (mW)	P_{SAVING} (mW
Conventional VRS	Temporary	97.88				236.34		334.22	–
	Permanent					65.83	5.72	71.55	–
LWC VRS	Temporary		10.89			83.99		94.88	239.34
	Permanent	97.88		15.36		189.54	5.57	308.35	–236.8
LWC + DS VRS	Temporary		10.89			81.07		91.96	242.26
	Permanent	24.47		15.36	13.64	107.4	4.71	165.58	–94.03
Low-power H.264 VRS	Temporary	60.94				165.23		226.17	108.05
	Permanent					65.88	6.72	72.6	–1.05

Simulation results (3)

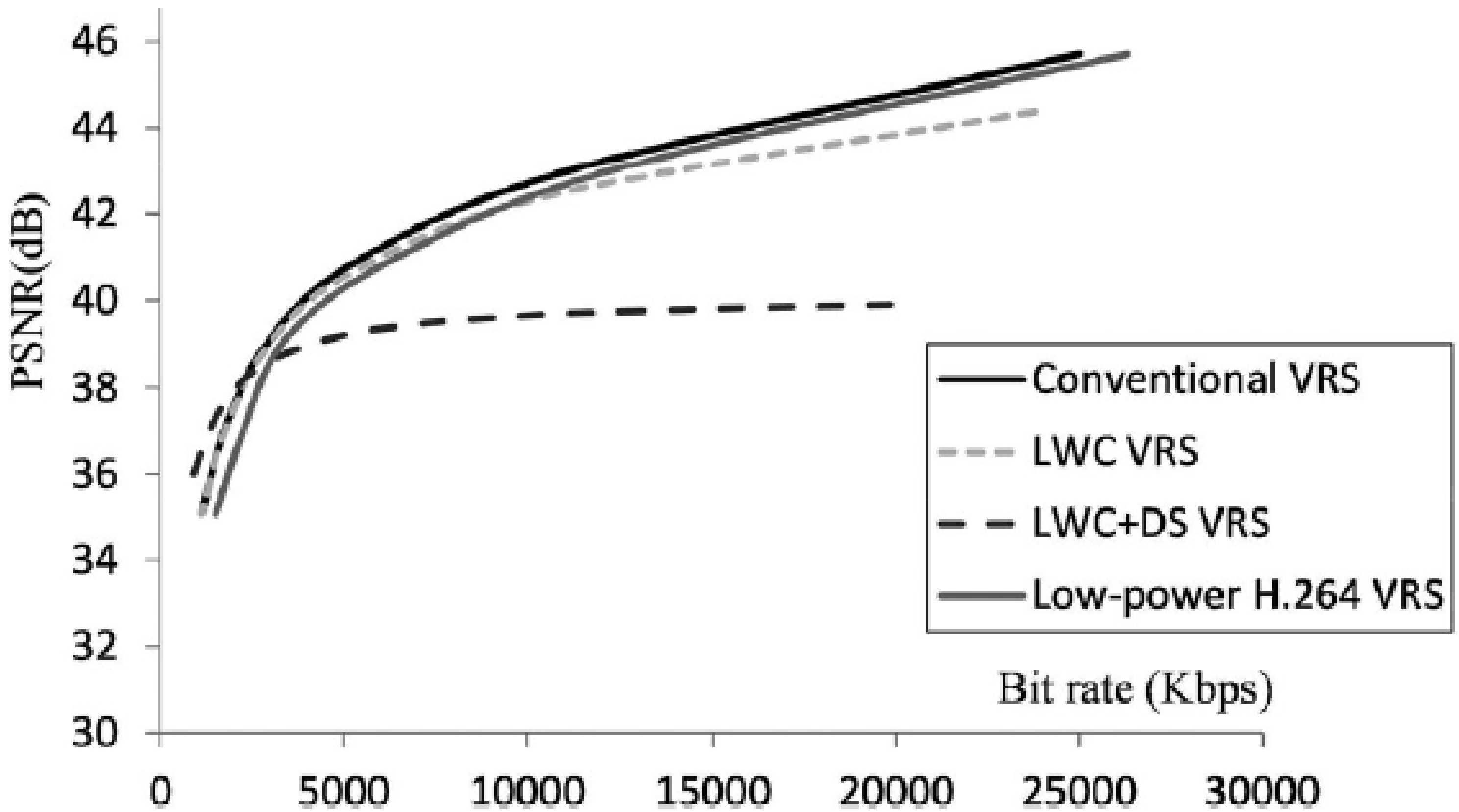
- Effect of the frequency of the permanent mode

POWER CHANGES ACCORDING TO CHANGES IN F_{RECORD}

F_{RECORD} (%)	Power Consumption (mW)				Power Gain (%)		
	Conventional	LWC	LWC + DS	LP H.264	LWC	LWC + DS	LP H.264
0	334.2	94.9	92.0	226.2	71.6	72.5	32.3
5	337.8	110.3	100.2	229.8	67.3	70.3	32.0
10	341.4	125.7	108.5	233.4	63.2	68.2	31.6
20	348.5	156.6	125.1	240.7	55.1	64.1	30.9
25	352.1	172.0	133.4	244.3	51.2	62.1	30.6
30	355.7	187.4	141.6	248.0	47.3	60.2	30.3
40	362.8	218.2	158.2	255.2	39.9	56.4	29.7
50	370.0	249.1	174.8	262.5	32.7	52.8	29.1
60	377.2	279.9	191.3	269.7	25.8	49.3	28.5
70	384.3	310.7	207.9	277.0	19.1	45.9	27.9
75	387.9	326.1	216.1	280.6	15.9	44.3	27.7
80	391.5	341.6	224.4	284.3	12.7	42.7	27.4
90	398.6	372.4	241.0	291.5	6.6	39.5	26.9
100	405.8	403.2	257.5	298.8	0.6	36.5	26.4

Simulation results (4)

- Compression efficiency



Simulation results (5)

RELATIONSHIP BETWEEN THE BDPSNR AND THE POWER SAVING WHEN F_{RECORD} IS 10%

F_{RECORD} = 10%	Low Target Bitrate			High Target Bitrate		
	LWC	LWC + DS	LP H.264	LWC	LWC + DS	LP H.264
BDPSNR(dB)	−0.031	0.07	−0.313	−0.323	−2.202	−0.165
PS(%)	63.2	68.2	31.6	63.2	68.2	31.6
BDPSNR/PS	−0.049	0.103	−0.991	−0.511	−3.229	−0.522

TABLE VII
RELATIONSHIP BETWEEN THE BDPSNR AND THE POWER SAVING WHEN F_{RECORD} IS 70%

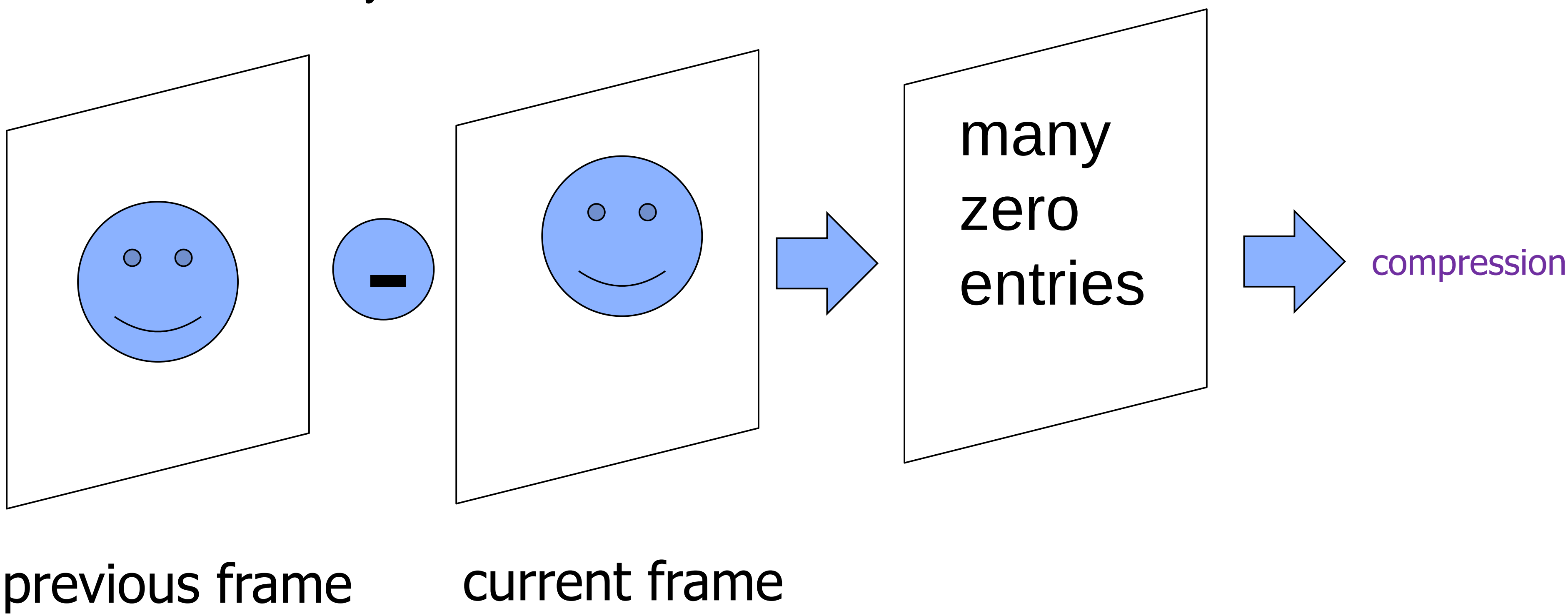
F_{RECORD} = 70%	Low Target Bitrate			High Target Bitrate		
	LWC	LWC + DS	LP H.264	LWC	LWC + DS	LP H.264
BDPSNR(dB)	−0.031	0.07	−0.313	−0.323	−2.202	−0.165
PS(%)	19.1	54.9	27.9	19.1	54.9	27.9
BDPSNR/PS	−0.162	0.128	−1.122	−1.691	−4.011	−0.591

Background subtraction w/ compression

- Background subtraction:
  - an important step for object detection
  - Reduce complexity/data in the background
- Avoid the complexity for background subtraction
  - Use the information obtained from video compression

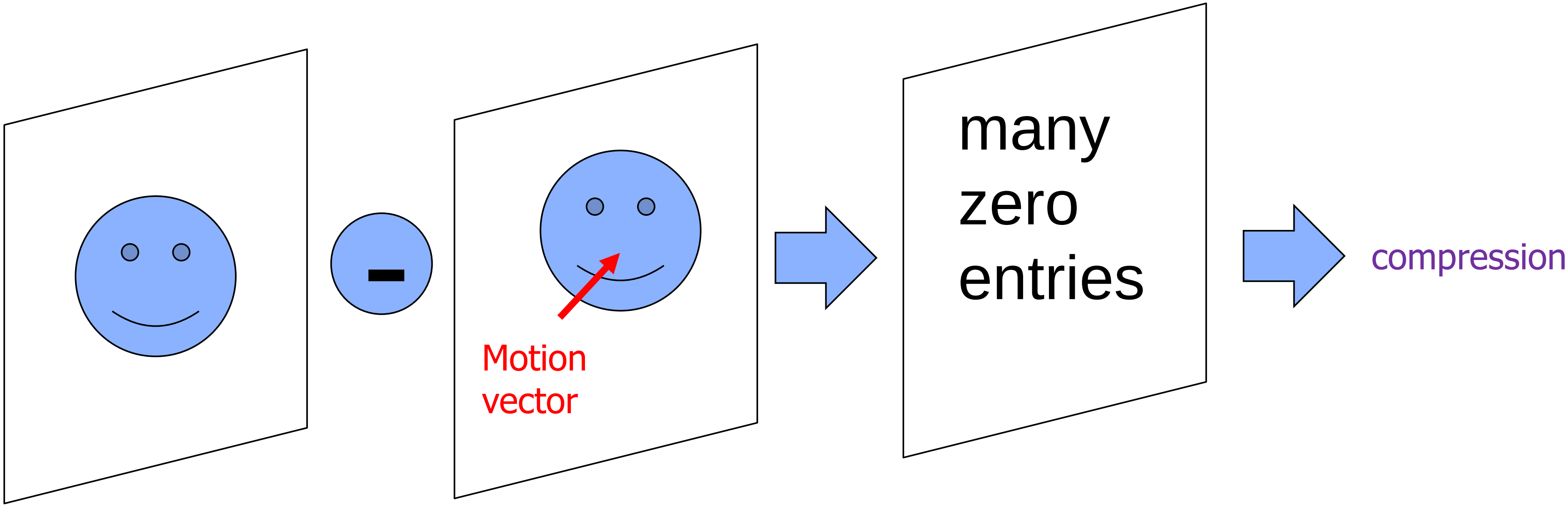
Information from Video Compression

- To remove the temporal redundancy,
  - (current frame – previous frame) is obtained
    - Many ‘0’s in this subtracted frame



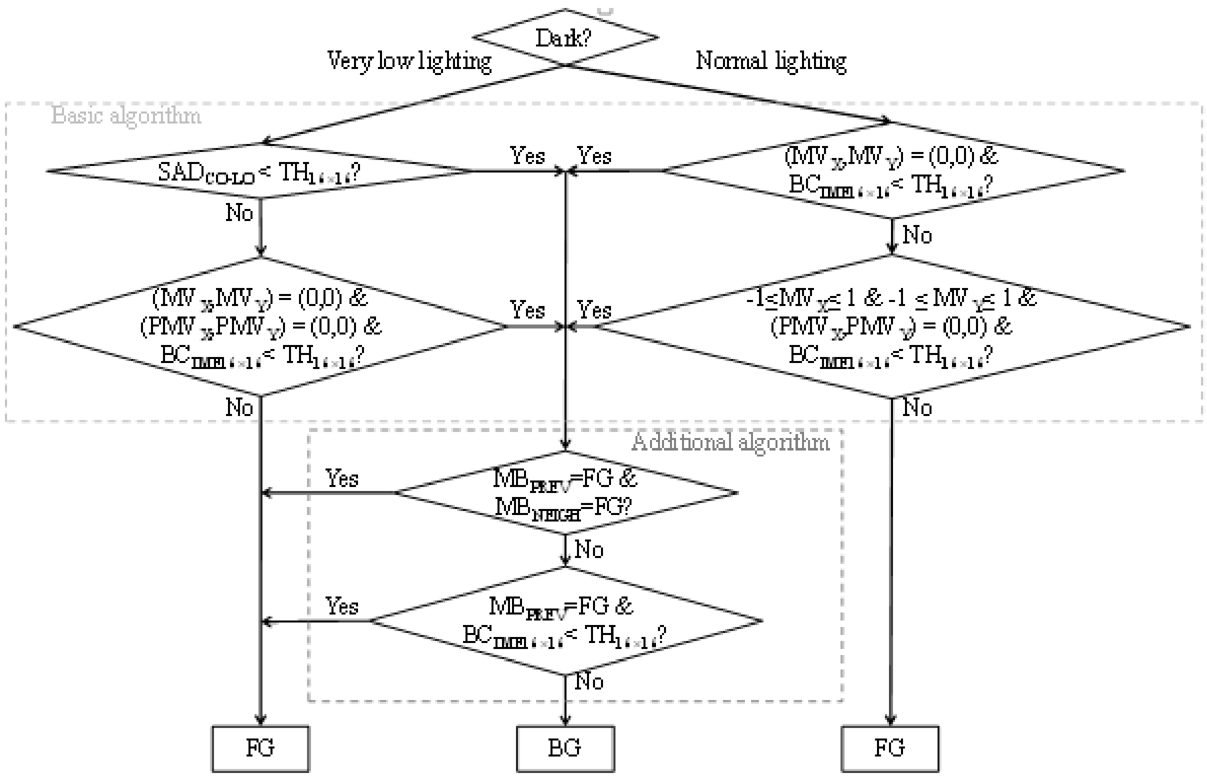
Information from Video Compression

- To make more zeros in the subtracted frame,
  - Find the best matching position
    - This is called “motion vector”



- The motion vector and matching errors are used to detect the background
  - Background: no motion and little matching error

Algorithm and Results



THE RATIOS OF FALSE-NEGATIVE AND FALSE-POSITIVE ERRORS

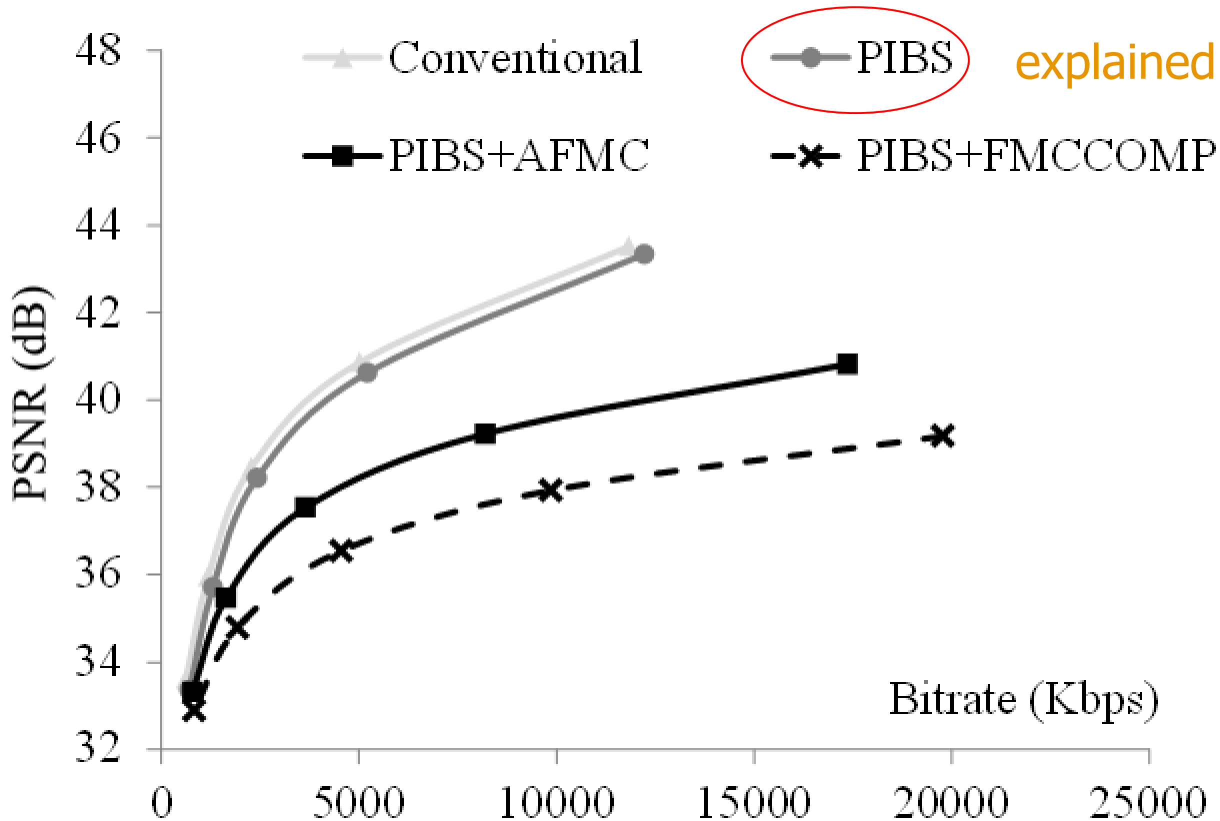
False-negative (%)	False-positive (%)
0.981	0.48

Video Coding Complexity Adjustment

- Background area:
  - May sacrifice video quality
  - Less efficiency compression algorithm
  - Less complex compression algorithm
- Coding options: Skip some coding options
  - FME (Fractional Motion Estimation): ¼ pixel precision for motion information derivation
  - IP (Intra-frame Prediction): removal of spatial redundancy

Classification	Coding option for the MB_{CURR}
Strong FG (M3)	Regular (FME & IP)
Object Boundary (M2)	Only IP
Uncovered BG (M1)	Only IP
Strong BG (M0)	SKIP mode (Skip both FME and IP)

Compression Efficiency



Contents

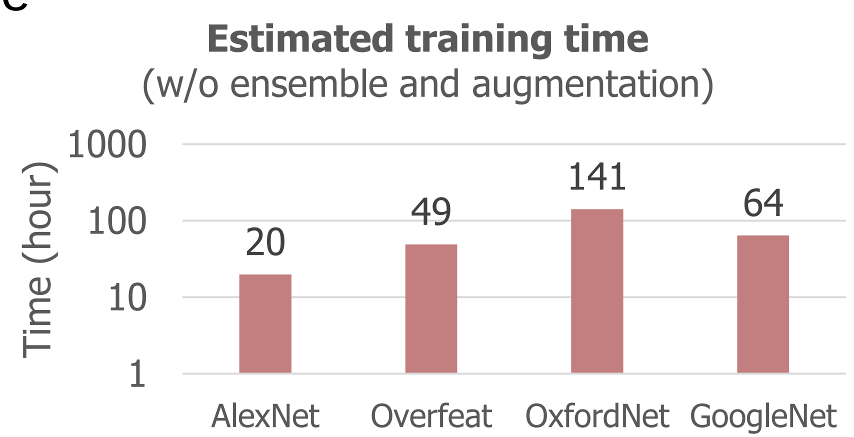
- Intelligent Video Cameras
- Power consumption in video cameras
- Reduction of complexity/power consumption
- Future trends

Data Explosion

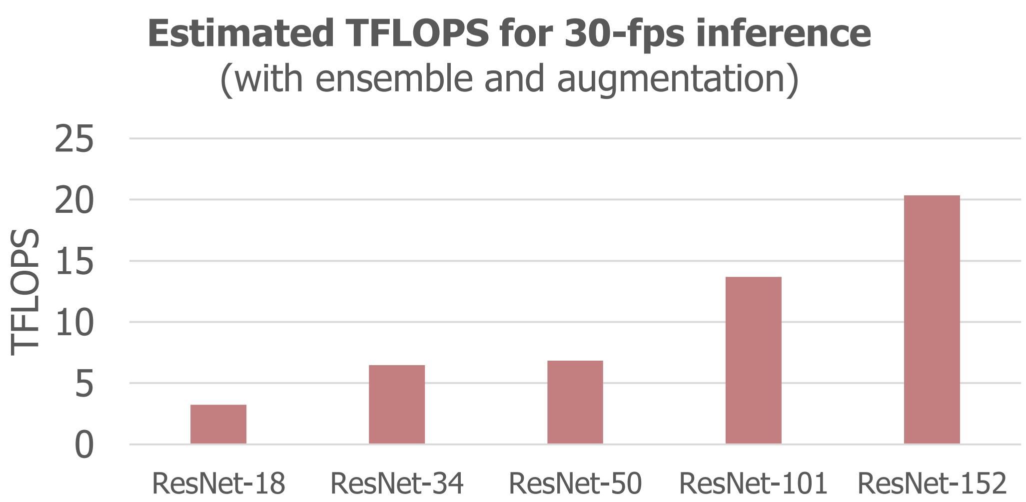
(source) 2016, IDF2016 keynote

Deep Learning Complexity

- Training Time



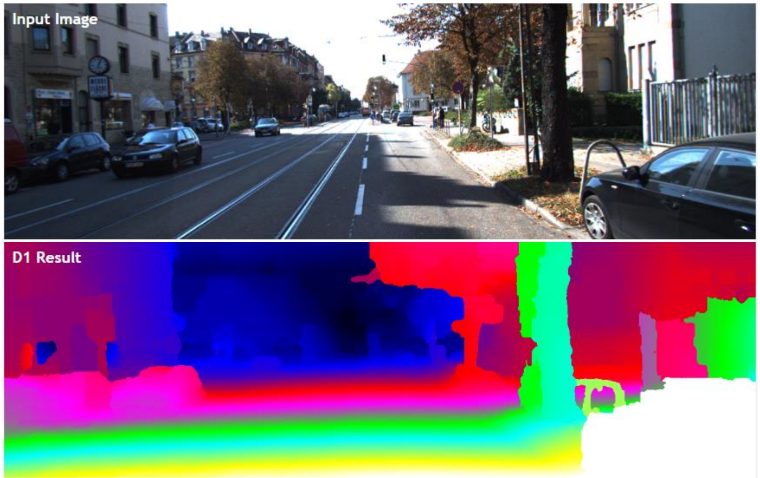
- Inference Time

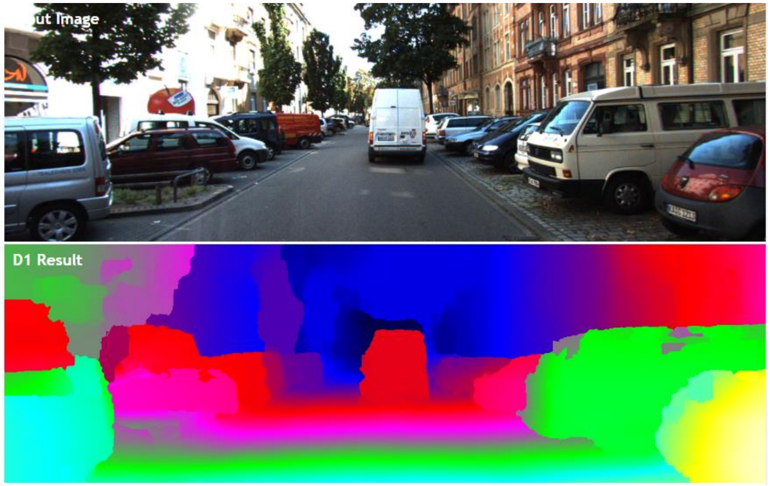


37

Distance generation complexity

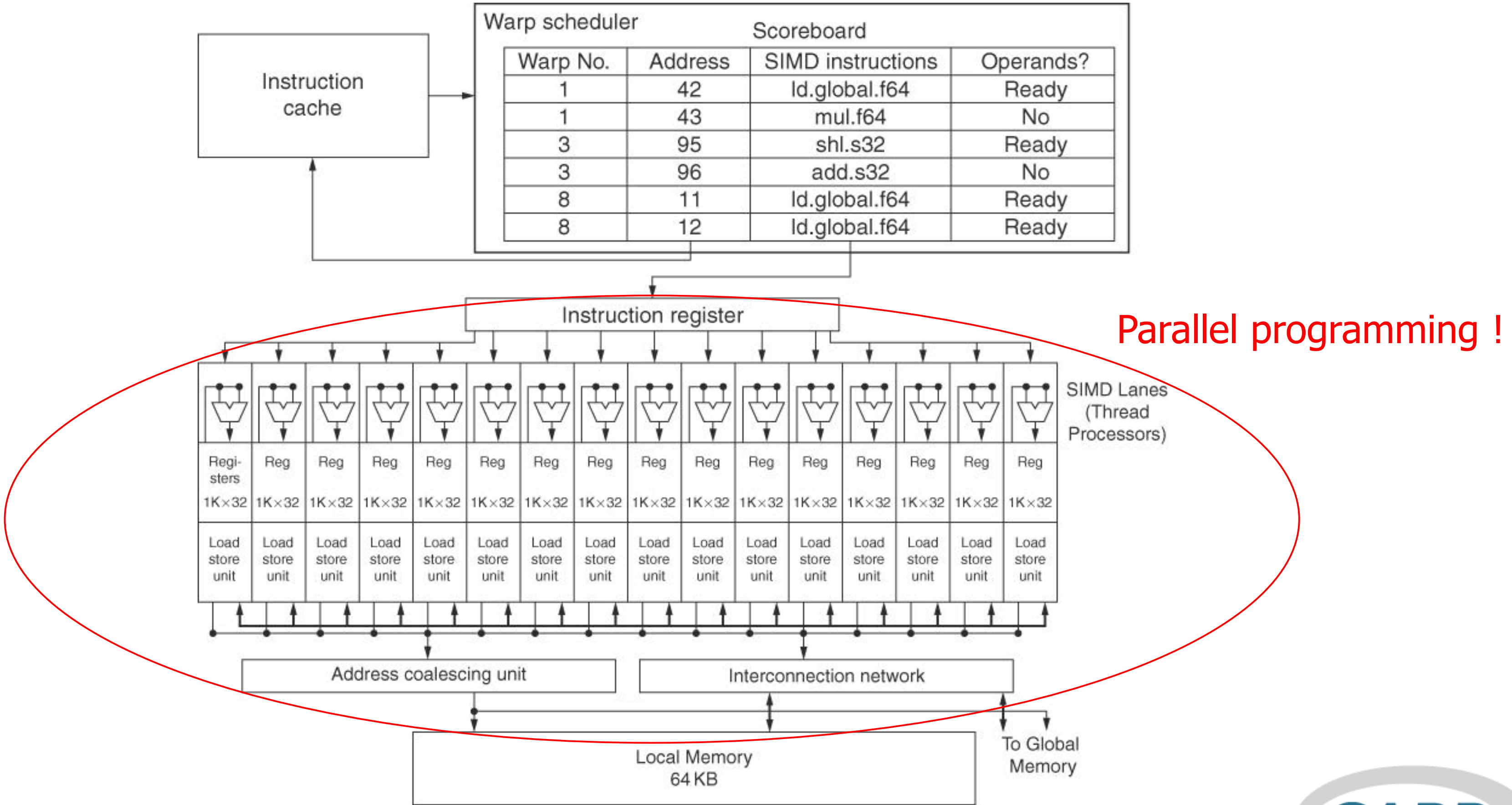
	Method	Data	Code	D1-bg	D1-fg	D1-all	Density	Time	Environment
1	Displets v2		code	3.00 %	5.56 %	3.43 %	100.00 %	265 s	>8 cores @ 3.0 Ghz (Matlab + C/C++)
F. Guney and A. Geiger: Displets: Resolving Stereo Ambiguities using Object Knowledge. Conference on Computer Vision and Pattern Recognition (CVPR) 2015.									
2	CNNF+SGM			2.78 %	7.69 %	3.60 %	100.00 %	71 s	GPU @ 2.5 Ghz (Python)
Anonymous submission									
3	PBCP			2.58 %	8.74 %	3.61 %	100.00 %	68 s	Nvidia GTX Titan X
A. Seki and M. Pollefeys: Patch Based Confidence Prediction for Dense Disparity Map. British Machine Vision Conference (BMVC) 2016.									
4	MC-CNN-acrt		code	2.89 %	8.88 %	3.89 %	100.00 %	67 s	Nvidia GTX Titan X (CUDA, Lua/Torch7)
J. Zbontar and Y. LeCun: Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches. Submitted to JMLR .									





Architecture complexity

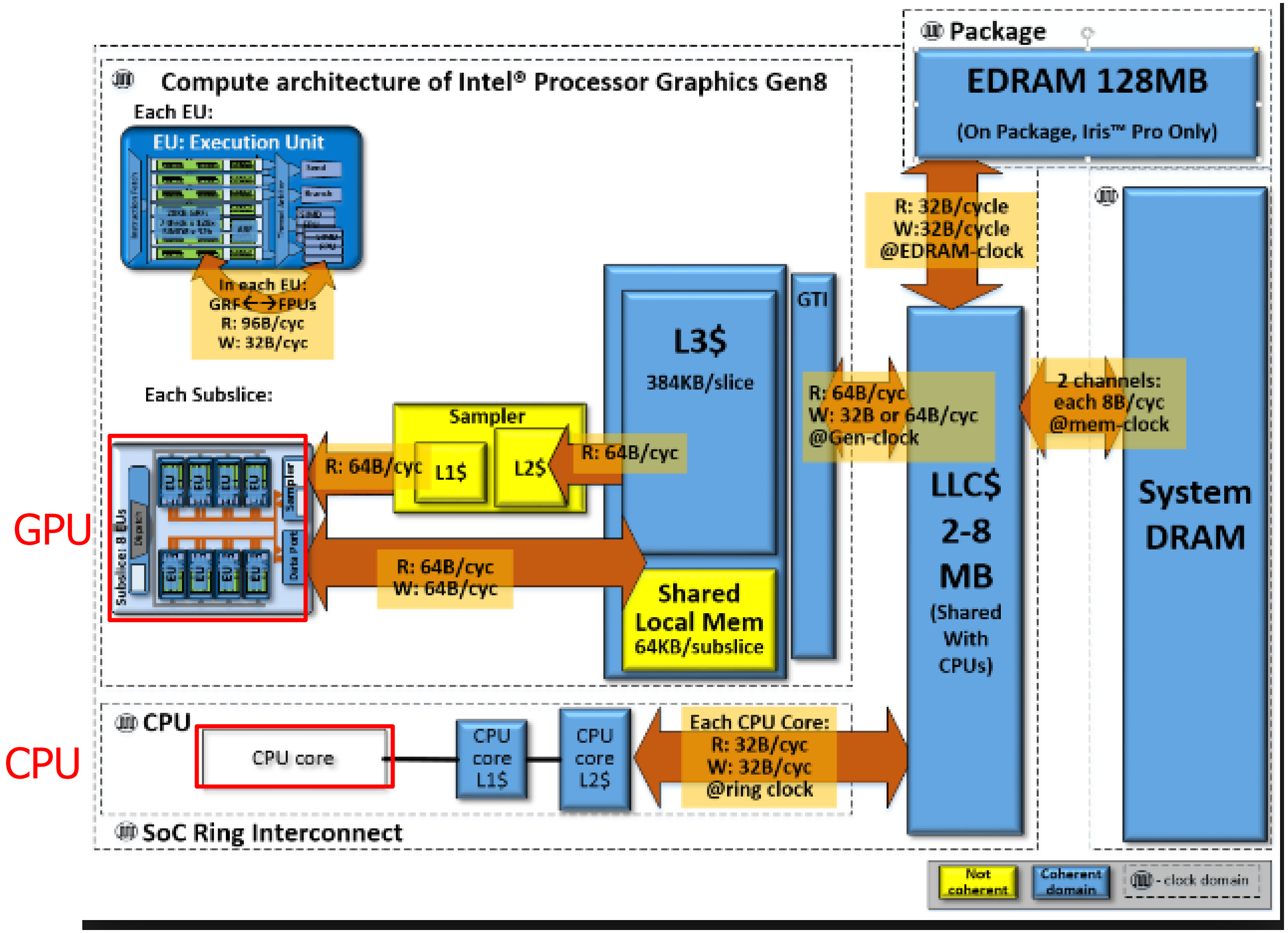
- GPU can easily make 10x speed-up
- However, you need to understand its architecture



Warp No.	Address	SIMD instructions	Operands?
1	42	ld.global.f64	Ready
1	43	mul.f64	No
3	95	shl.s32	Ready
3	96	add.s32	No
8	11	ld.global.f64	Ready
8	12	ld.global.f64	Ready

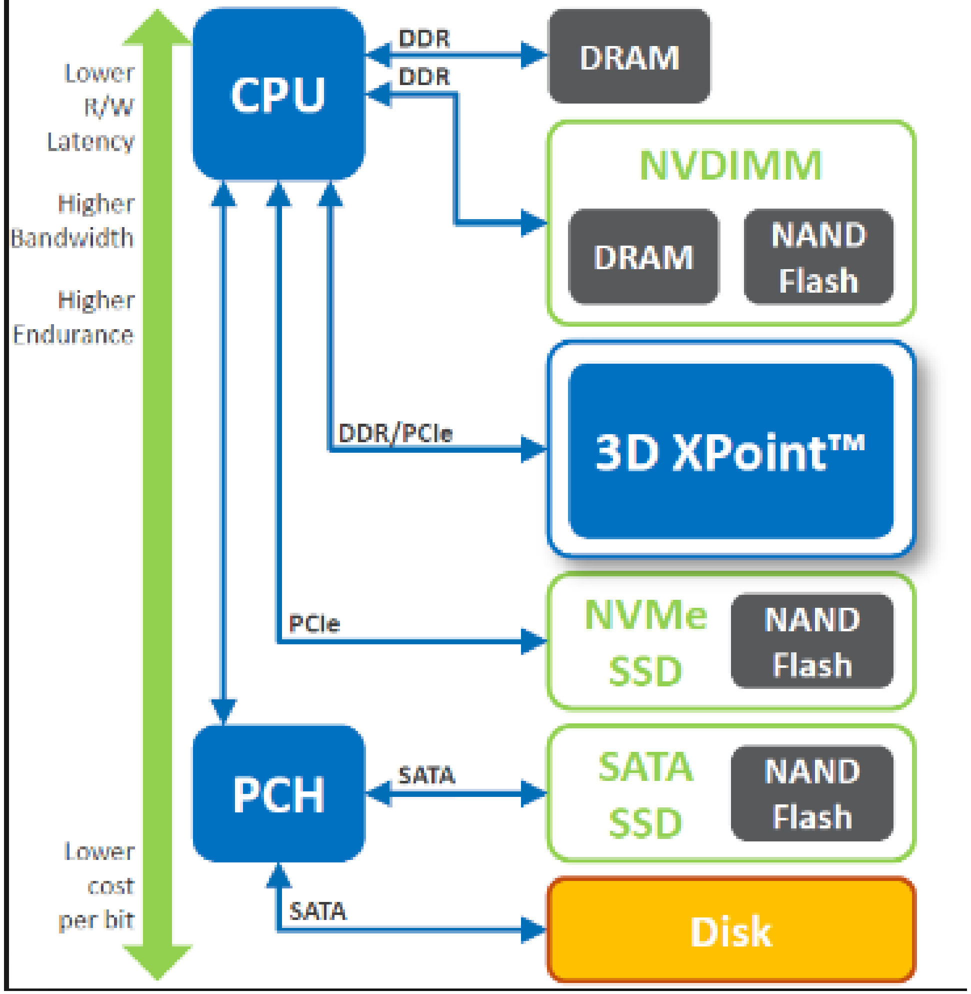
Intel CPU+GPU Architecture

- Intel Gen8 Architecture:
  - Complex memory architecture to share data between CPU & GPU



Complex memory hierarchy outside CPU

- Server architecture



Conclusions

- Huge amount of data processing
  - Need to reduce the amount of data to be stored or transmitted
  - Requires complex computation to find redundancy
    - Trade off between data amount and computation complexity
    - Need to reduce the computation complexity
- Complex computer architecture for fast data processing

  → Efficient use of complex architecture is important
- Video compression is a good example

Appendix

Global Research Network Program

- Support research activities for international collaboration
  - KRW: USD ~ 1000:1

	2015	2016
Total fund	6.6 billion KRW (~6.6 million USD)	75.9 billion KRW
# of new projects	1 (89 million KRW)	26 (3, 115 million KRW) 119 million KRW/project
period	1 ~ 3 years	2 ~ 3 years
subject	Korea-china network	Social science/liberal art

- Additional programs for one-to-one collaboration