From image coding and representation to robotic vision

Marie Babel

Lagadic group
INRIA Rennes Bretagne-Atlantique & IRISA
http://www.irisa.fr/lagadic

Habilitation thesis - 2012-06-29

Outline

1. Professional background
2. Research activities
3. Coding and services for image and video transmission
4. Representation frameworks
5. Research project: towards robotic vision

Outline

Professional background

- 2002: Electronic and Applied Computer Science Engineer
- 2002-2005: PhD on image processing
  - INSA de Rennes, IETR Lab
  - *Lossless and lossy image compression through the Locally Adaptive Resolution (LAR) method*
  - Supervisors: Joseph Ronsin - Olivier Déforges
- 2005-2006: ATER - INSA de Rennes
- 2006: Assistant professor - INSA de Rennes
  - *Teaching*: Department of Electronic and Applied Computer Science
  - *Research*: IETR Lab, Image group
- 2011: IRISA / INRIA Lab, Lagadic Team

Administrative and teaching activities

- Administrative activities
  - 1999-2000: *Conseil d'Administration* - INSA Rennes
  - 2010-2011: *Conseil de Composante* - IETR Rennes
  - 2010-2012: *Conseil Scientifique* - INSA Rennes
  - 2012: *Comité de Centre* - INRIA Rennes
- Teaching activities
  - INSA Rennes
    - Electronic and Applied Computer Science department - Master level
    - Master of Science: i-MARS
    - Computer Science department - Master level
  - Signal and image processing
  - Electronics

Outline

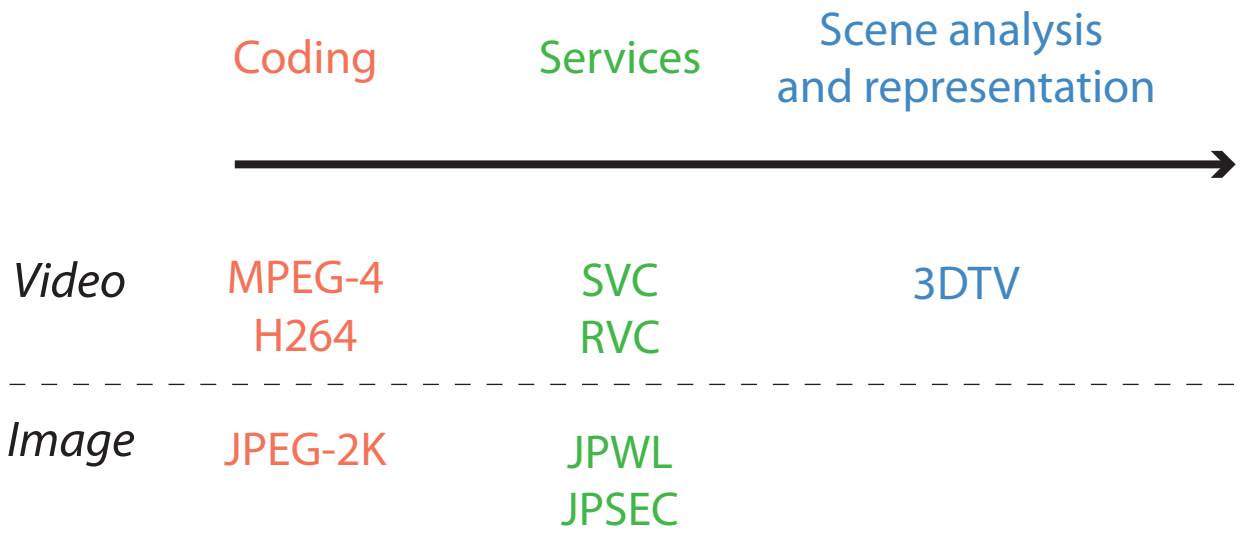
Research activities: overview

- Progressive evolution of research topics

- Key words
  - Compression and coding
  - Quality of Service / Quality of Experience
  - Pseudo-semantic temporal analysis
- Towards robotic vision: related research topics
  - Region/object representation
  - 3D tracking
  - Object motion and trajectory analysis

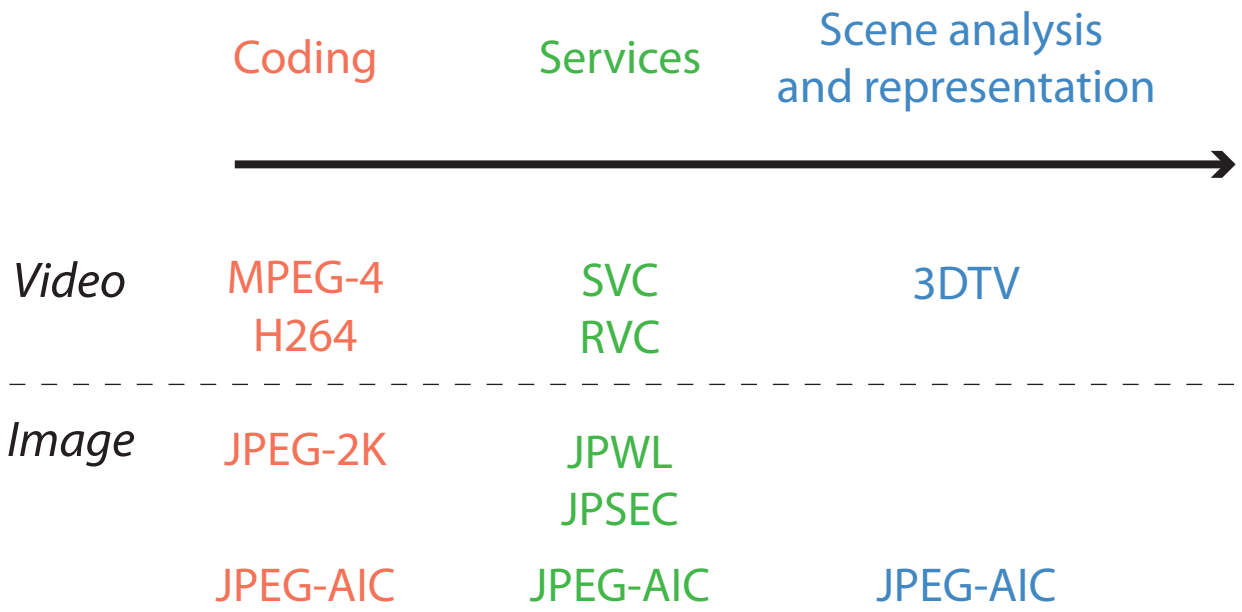
Research activities: overview

- International activity context: standards



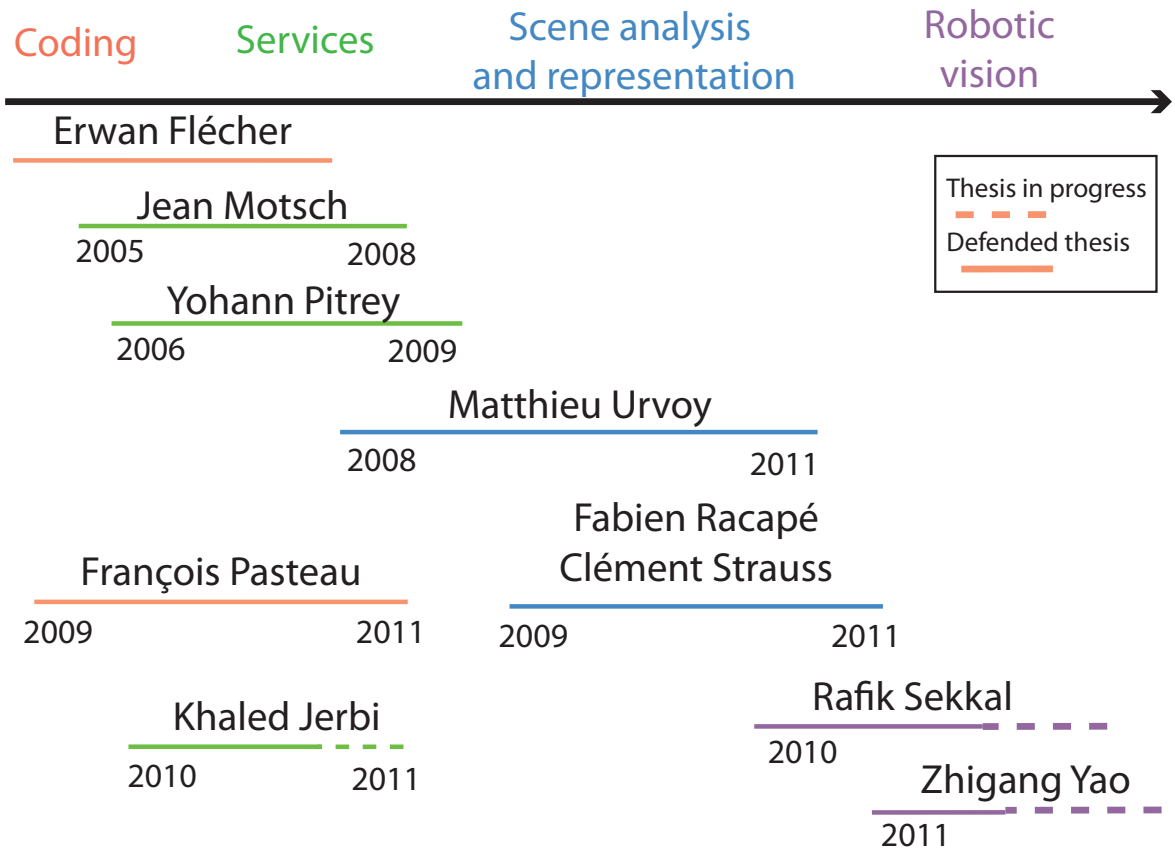
Research activities: overview

- International activity context: standards



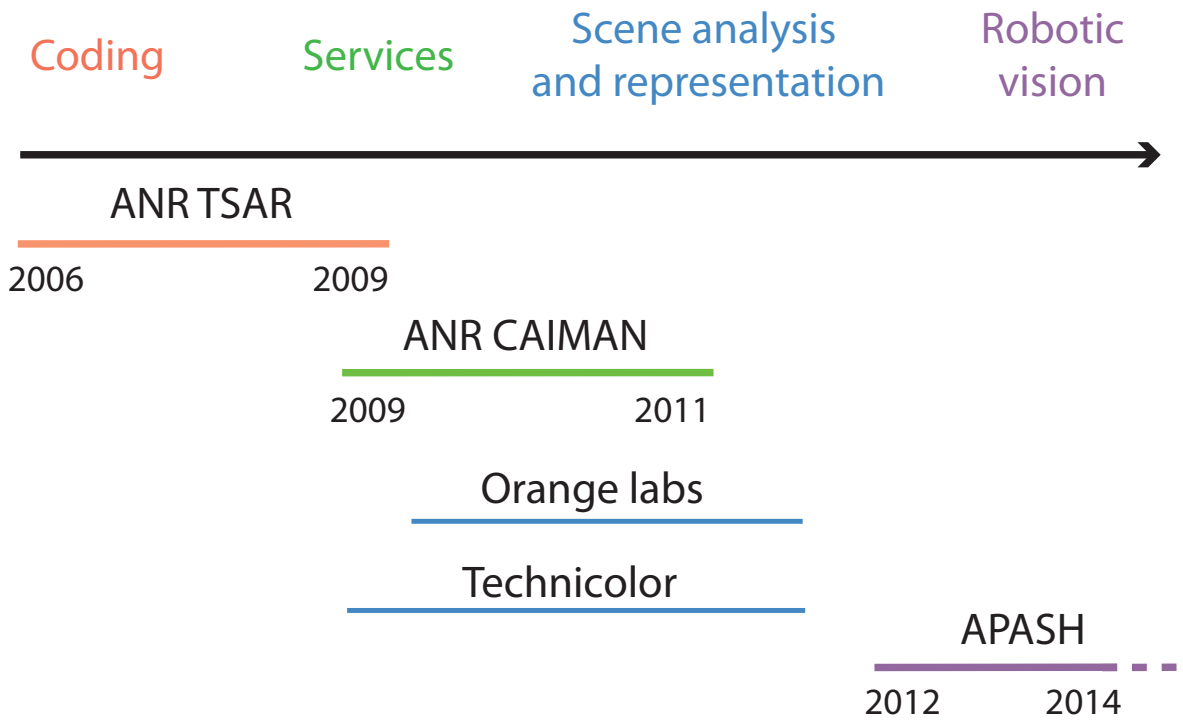
Research activities: overview

- Co-supervision of 10 PhD students



Research activities: overview

- Collaborative projects



- CAIMAN - TSAR: scientific leader of IETR
- APASH: project coordinator

Outline

Coding research context

- IETR Laboratory
  - Locally Adaptive Resolution (LAR) coding scheme
  - Low complex codec designed for embedded systems
- ANR TSAR (Transmission Sécurisée d'images d'Art haute Résolution)
- Joint ANR CAIMAN (Codeur Avancé d'IMAges et Nouveaux services) and JPEG Advanced Image Coding
  - New coding framework associated to advanced services
  - Generic Quality of Service
  - Generic Quality of Experience

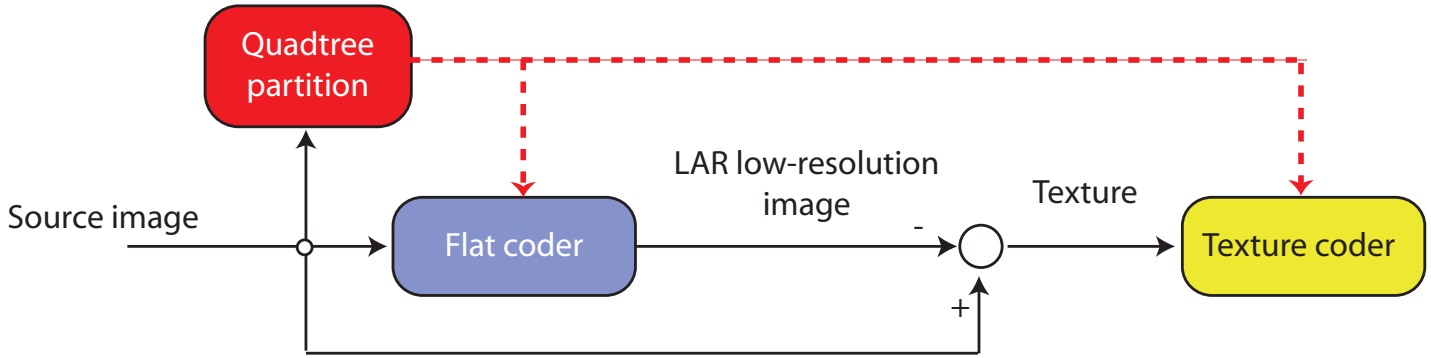
Coding research context

- IETR Laboratory
  - Locally Adaptive Resolution (LAR) coding scheme
  - Low complex codec designed for embedded systems
- ANR TSAR (Transmission Sécurisée d'images d'Art haute Résolution)
- Joint ANR CAIMAN (Codeur Avancé d'IMAges et Nouveaux services) and JPEG Advanced Image Coding
  - New coding framework associated to advanced services
  - Generic Quality of Service
  - Generic Quality of Experience

$\rightarrow$ LAR: response to JPEG-AIC call for proposal

LAR framework

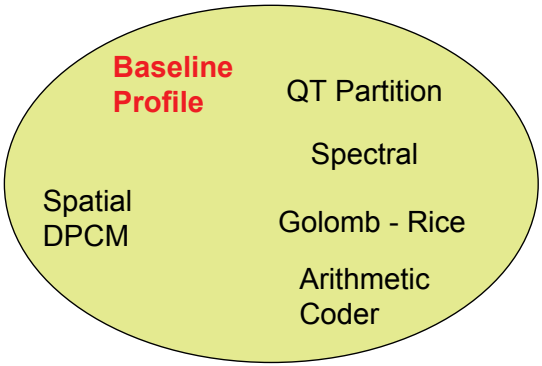
- Joint representation and coding issues
  - Two-layer codec
  - Quadtree based framework



- Contributions
  - Pyramidal solution for lossless compression
  - Data integrity
  - Robustness

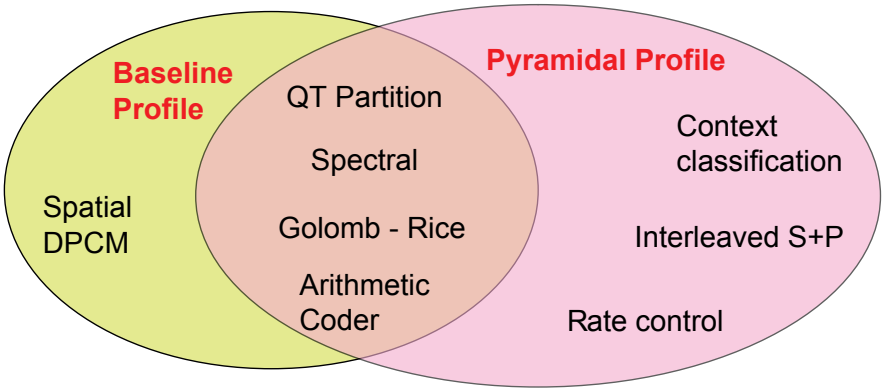
LAR Profiles

- Controlled complexity: 3 profiles
  - Baseline profile



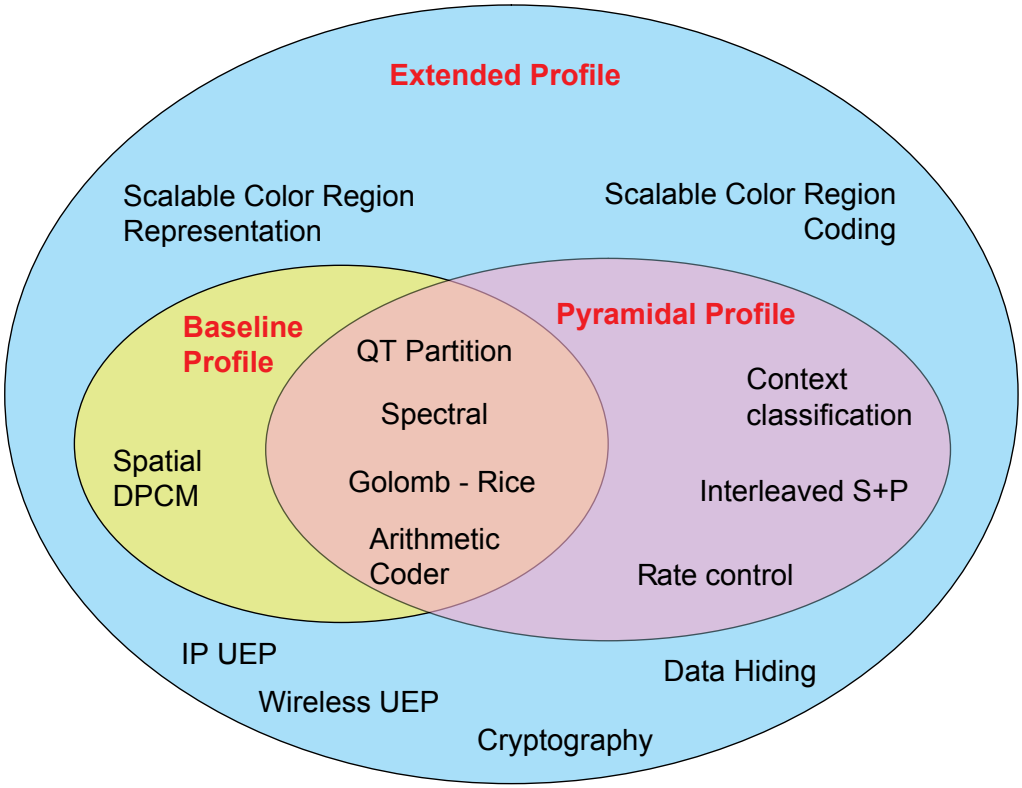
LAR Profiles

- Controlled complexity: 3 profiles
  - Pyramidal profile [Babel05][Pasteau11]: Interleaved S+P



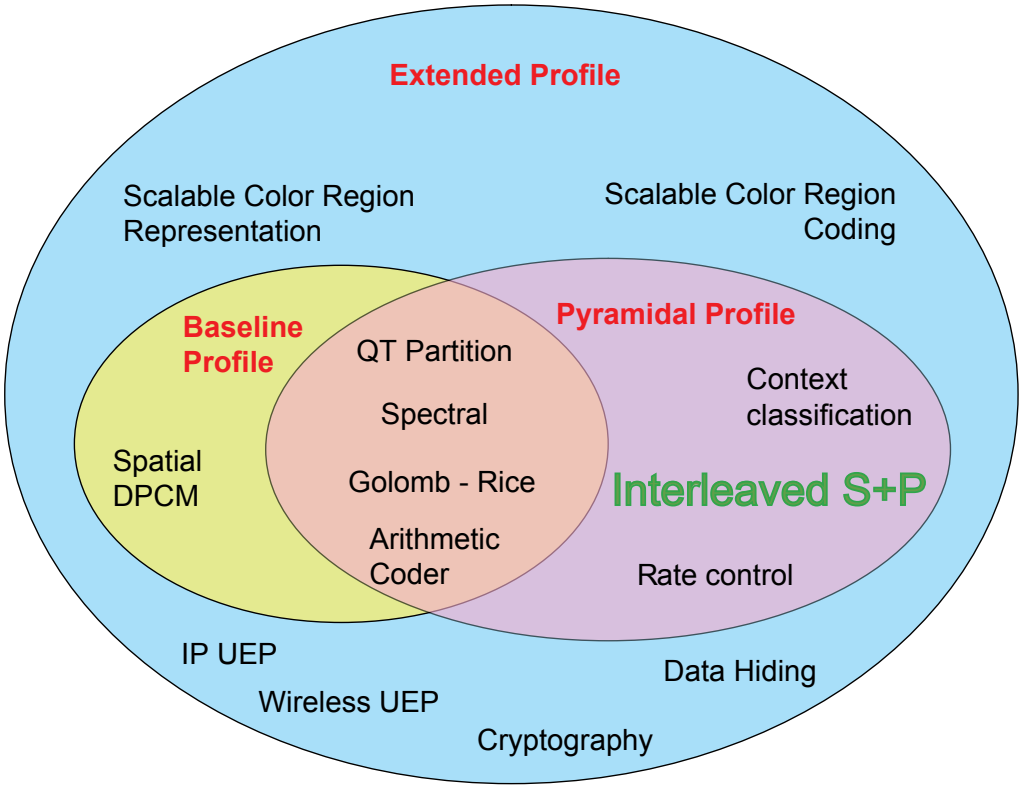
LAR Profiles

- Controlled complexity: 3 profiles
  - Extended profile [Flecher07][Motsch08][Strauss11]



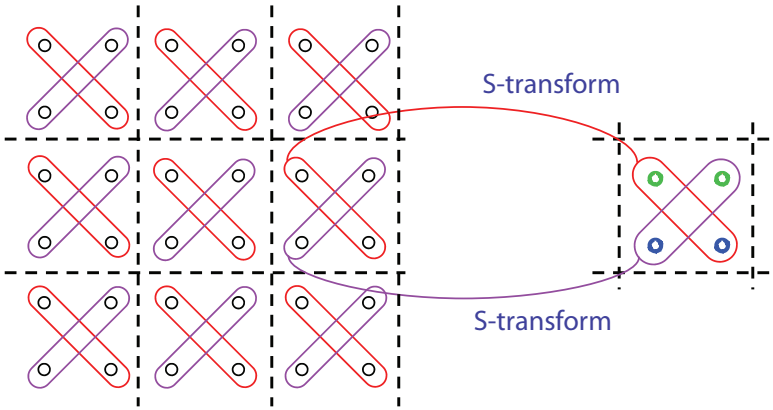
LAR Profiles

- Controlled complexity: 3 profiles
  - Extended profile [Flecher07][Motsch08][Strauss11]



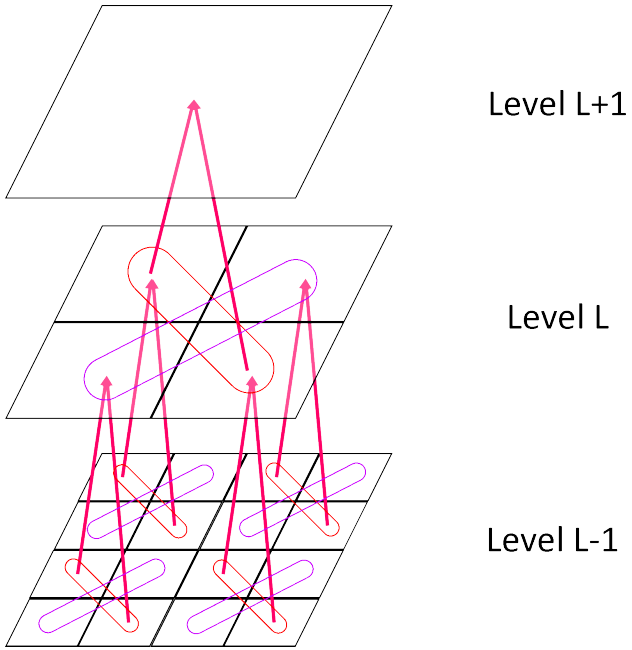
Interleaved S+P framework

- Principles [BDR05]
  - Two Interleaved S-Transforms within 2x2 blocks



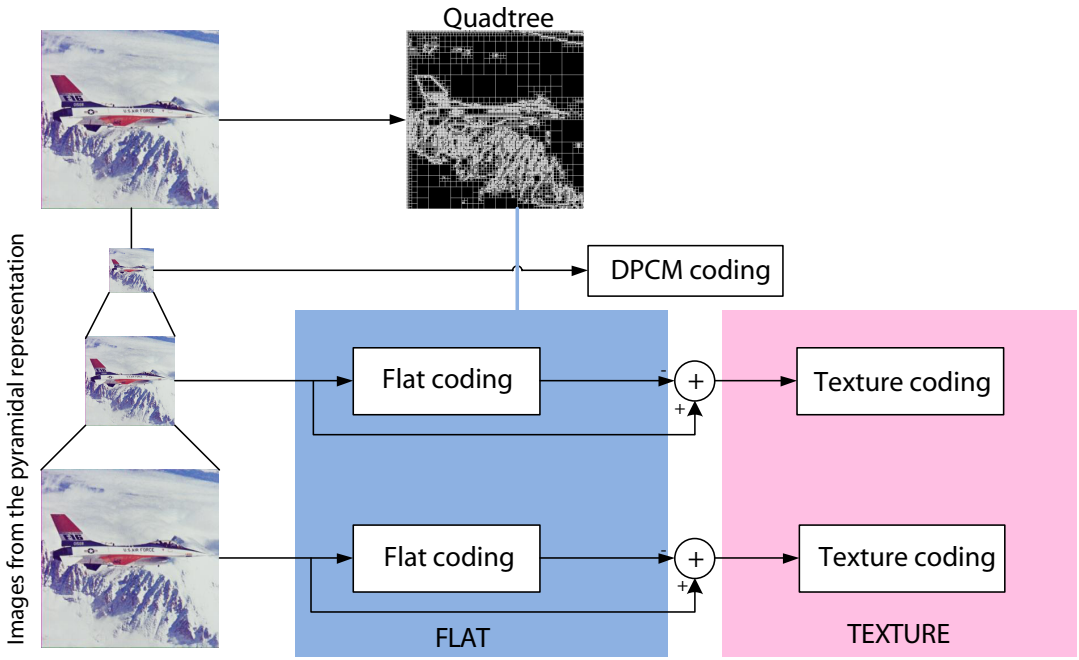
Interleaved S+P framework

- Principles [BDR05]
  - Two Interleaved S-Transforms within 2x2 blocks
  - Extension to multiresolution



Interleaved S+P framework

- Principles [BDR05]
  - Two Interleaved S-Transforms within 2x2 blocks
  - Extension to multiresolution
  - Pyramidal decomposition with predictive schemes



Interleaved S+P framework

- Properties
  - Scalable
  - Efficient lossless compression
    - medical images [BPSPBBD11]: rate decrease of 20% on mammograms
  - Intrinsic robustness to localized reconstruction errors
  - Efficient hardware implementations - Embedded systems
    - FPGA implementation [DB08]
    - Co-simulation and rapid prototyping [FRRBD07]
    - RVC framework [JWRDBA11]

Interleaved S+P framework

- Properties
  - Scalable
  - Efficient lossless compression
    - medical images [BPSPBBD11]: rate decrease of 20% on mammograms
  - Intrinsic robustness to localized reconstruction errors
  - Efficient hardware implementations - Embedded systems
    - FPGA implementation [DB08]
    - Co-simulation and rapid prototyping [FRRBD07]
    - RVC framework [JWRDBA11]

$\rightarrow$ Properties required for JPEG-AIC

JPEG-AIC standardization

- Responses to Call for Technologies
  - Natural and medical image coding tools
- Interleaved S+P
  - 4 core experiments
  - 10 contributions (2008-2011)
- Results
  - Objective and subjective quality in between JPEG and JPEG-XR
  - Complexity similar to JPEG2K
  - Academic solution: non-optimal implementation
  - Standardization process stopped in 2011

JPEG-AIC standardization

- Responses to Call for Technologies
  - Natural and medical image coding tools
- Interleaved S+P
  - 4 core experiments
  - 10 contributions (2008-2011)
- Results
  - Objective and subjective quality in between JPEG and JPEG-XR
  - Complexity similar to JPEG2K
  - Academic solution: non-optimal implementation
  - Standardization process stopped in 2011

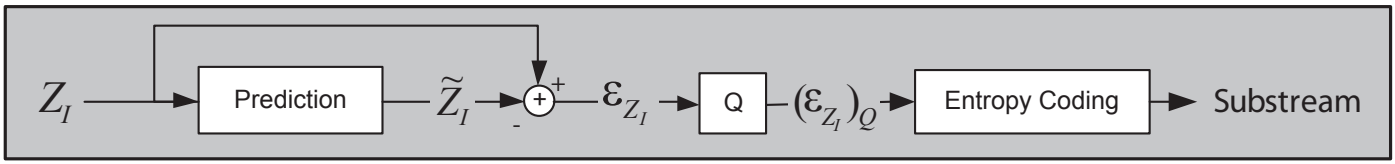
$\rightarrow$ Efficient generic coding solutions

Generic predictive coding tools

- Objectives: QoS
  - Defining generic Rate/Distorsion models for predictive coders
  - Designing low complex solutions

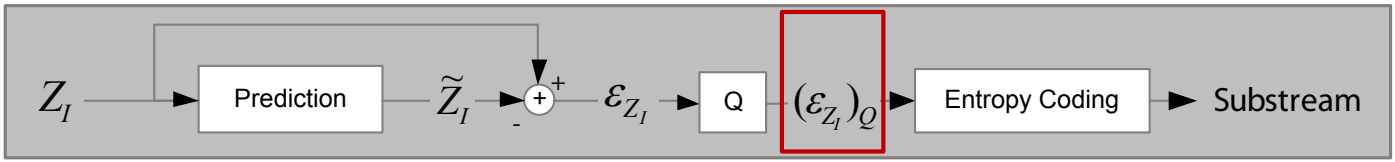
Generic predictive coding tools

- Objectives: QoS
  - Defining generic Rate/Distorsion models for predictive coders
  - Designing low complex solutions
- Functional objectives [Pasteau11]
  - Extraction of prediction error statistical models



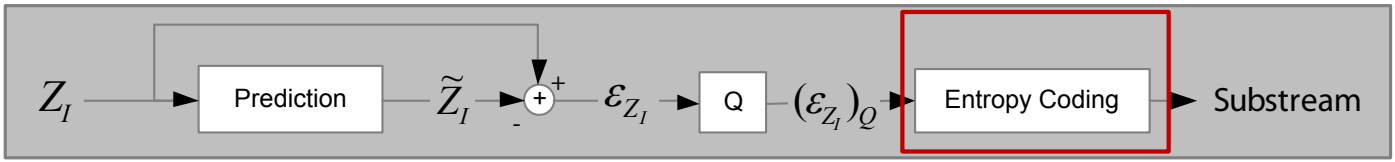
Generic predictive coding tools

- Objectives: QoS
  - Defining generic Rate/Distorsion models for predictive coders
  - Designing low complex solutions
- Functional objectives [Pasteau11]
  - Extraction of prediction error statistical models
  - Design of adaptive multi-component decorrelation



Generic predictive coding tools

- Objectives: QoS
  - Defining generic Rate/Distorsion models for predictive coders
  - Designing low complex solutions
- Functional objectives [Pasteau11]
  - Extraction of prediction error statistical models
  - Design of adaptive multi-component decorrelation
  - Low complex entropy coder



Generic predictive coding tools

- Prediction error statistical modeling
  - Hypothesis: Laplacian distribution
  - Principle: modeling of quantization effects
  - Results: Entropy and distortion modeling
    - Scalability: intra/inter level models

Generic predictive coding tools

- Prediction error statistical modeling
  - Hypothesis: Laplacian distribution
  - Principle: modeling of quantization effects
  - Results: Entropy and distortion modeling
    - Scalability: intra/inter level models
- Adaptive decorrelation during prediction process
  - Lossless color image compression: better than JPEG2K
  - Lossy coding: bitrate decrease / Objective quality increase

Generic predictive coding tools

- Prediction error statistical modeling
  - Hypothesis: Laplacian distribution
  - Principle: modeling of quantization effects
  - Results: Entropy and distortion modeling
    - Scalability: intra/inter level models
- Adaptive decorrelation during prediction process
  - Lossless color image compression: better than JPEG2K
  - Lossy coding: bitrate decrease / Objective quality increase
- Symbol-oriented QM coder
  - Based on Q15 standard: bit-plane arithmetic codec
  - Limited rate loss - 3%
  - 50% computational time saving

Generic predictive coding tools

- To sum up...
  - Predictive codec oriented solutions (JPEG-LS, CALIC, H264...)
  - Low complex
  - Coding efficiency enhancement
- Potential use cases
  - Rate / Distorsion optimization mechanism

Generic predictive coding tools

- To sum up...
  - Predictive codec oriented solutions (JPEG-LS, CALIC, H264...)
  - Low complex
  - Coding efficiency enhancement
- Potential use cases
  - Rate / Distorsion optimization mechanism
  - QoS / QoE

Generic service-oriented tools

- Sensitive data exchanges
  - Medical images: privacy, hidden metadata, streaming
  - Art images (TSAR project): copyright, end-user quality

- Contributions

Data protection	QoS / QoE
Steganography and cryptography [Motsch08]	One-pass rate control (SVC framework) [Pitrey09]
Unequal Error Protection [BPDNGC08][HOBDBL09]	Low complex interpolation method [Strauss11]

Generic service-oriented tools

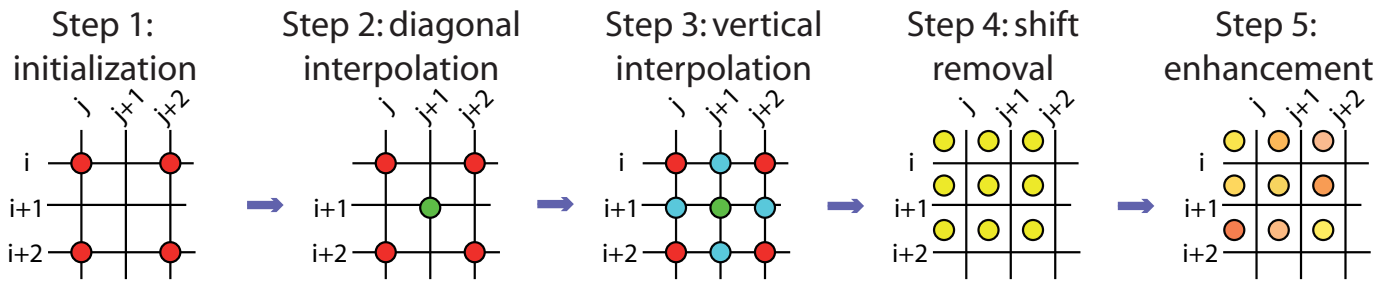
- Sensitive data exchanges
  - Medical images: privacy, hidden metadata, streaming
  - Art images (TSAR project): copyright, end-user quality

- Contributions

Data protection	QoS / QoE
Steganography and cryptography [Motsch08]	One-pass rate control (SVC framework) [Pitrey09]
Unequal Error Protection [BPDNGC08][HOBDBL09]	Low complex interpolation method [Strauss11]

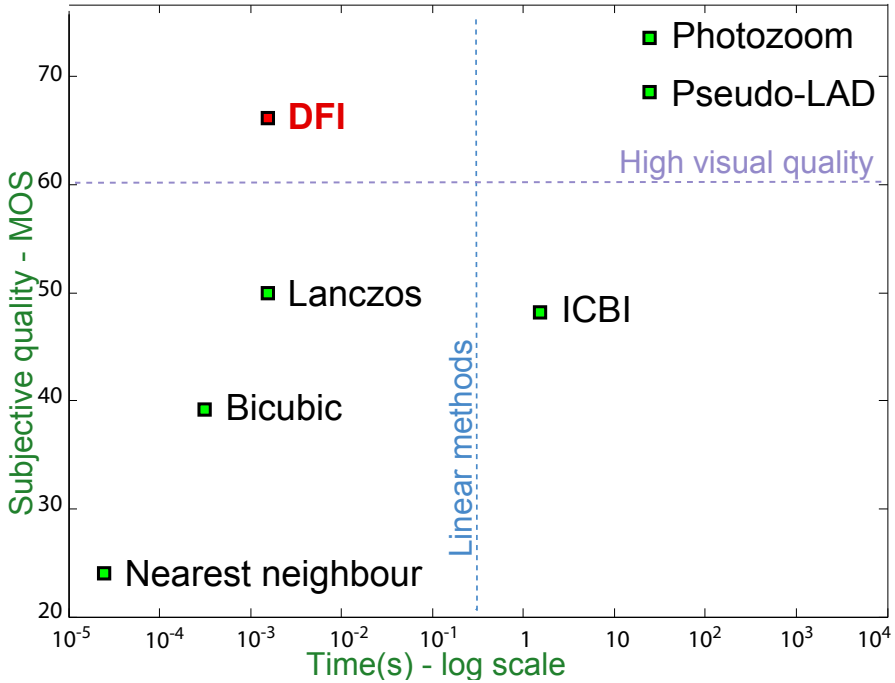
Dyadic Fast Interpolation (DFI)

- Motivations
  - Low complex interpolation method
  - High visual quality
- Principles
  - Five-step framework
  - Based on geometric duality principle
  - Only addition and shift operations



Dyadic Fast Interpolation

- Subjective assessments: SAMVIQ (VQEG) protocol



- Comparable quality with Photozoom and pseudo-LAD methods
- Complexity similar to linear methods
- Parallel implementation

Coding and services tools: conclusion

- Innovative coding frameworks
  - Scalable oriented tools
  - Generic tools for compression purposes
  - QoS / QoE services
  - Embedded systems concerns
- QoE next issues
  - Content-based solutions

Coding and services tools: conclusion

- Innovative coding frameworks
  - Scalable oriented tools
  - Generic tools for compression purposes
  - QoS / QoE services
  - Embedded systems concerns
- QoE next issues
  - Content-based solutions
  - Advanced image and video representations

Outline

1. Professional background
2. Research activities
3. Coding and services for image and video transmission
4. Representation frameworks
   - Quadtree-based segmentation
   - Motion tube representation
   - Adaptive image synthesis
5. Research project: towards robotic vision

Representation-oriented frameworks

- From pixel to region representation
  - Object-level representation
    - High level semantic data required
    - Shape, a priori behavior...
  - Intermediate representation: pseudo-semantic representation
    - Coherence in terms of texture, motion, color...
    - No side information
- Potential target applications
  - Coding
  - Texture / patch tracking
- Focus on representation aspects only

Representation-oriented frameworks

- Quadtree-based segmentation process
  - Quadtree structure: able to represent image content [Strauss11]
  - Region Adjacency Graph (RAG) based segmentation
  - Extension to scalability [Sekkal12]
  - Extension to video processing [Flecher08]

Representation-oriented frameworks

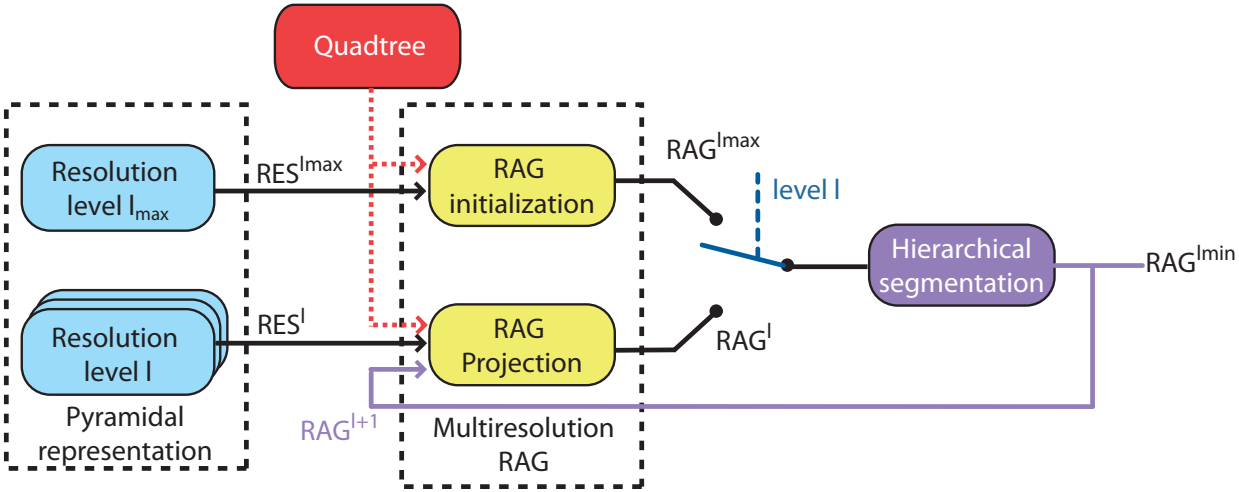
- Quadtree-based segmentation process
  - Quadtree structure: able to represent image content [Strauss11]
  - Region Adjacency Graph (RAG) based segmentation
  - Extension to scalability [Sekkal12]
  - Extension to video processing [Flecher08]
- Alternative representations for video coding standards
  - Rely on local cue and texture analysis
  - Block-based solutions
  - Motion tubes (Orange Labs - [Urvoy11])
  - Analysis/synthesis framework (Technicolor - [Racape11])

Spatial quadtree-based segmentation

- Spatial segmentation
  - Hierarchical split / merge framework
  - Color-based segmentation: joint analysis and coding [DBBR07]

Spatial quadtree-based segmentation

- Spatial segmentation
  - Hierarchical split / merge framework
  - Color-based segmentation: joint analysis and coding [DBBR07]
- Multiresolution region representation [SSPBD12]
  - Top-down process across resolution
  - RAG projection mechanism following quadtree structure

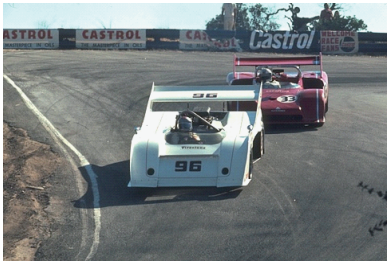


Spatial quadtree-based segmentation

- Properties
  - Low complexity: reduced number of blocks
  - Consistency via inherited labels
  - Robustness onto color gradated areas

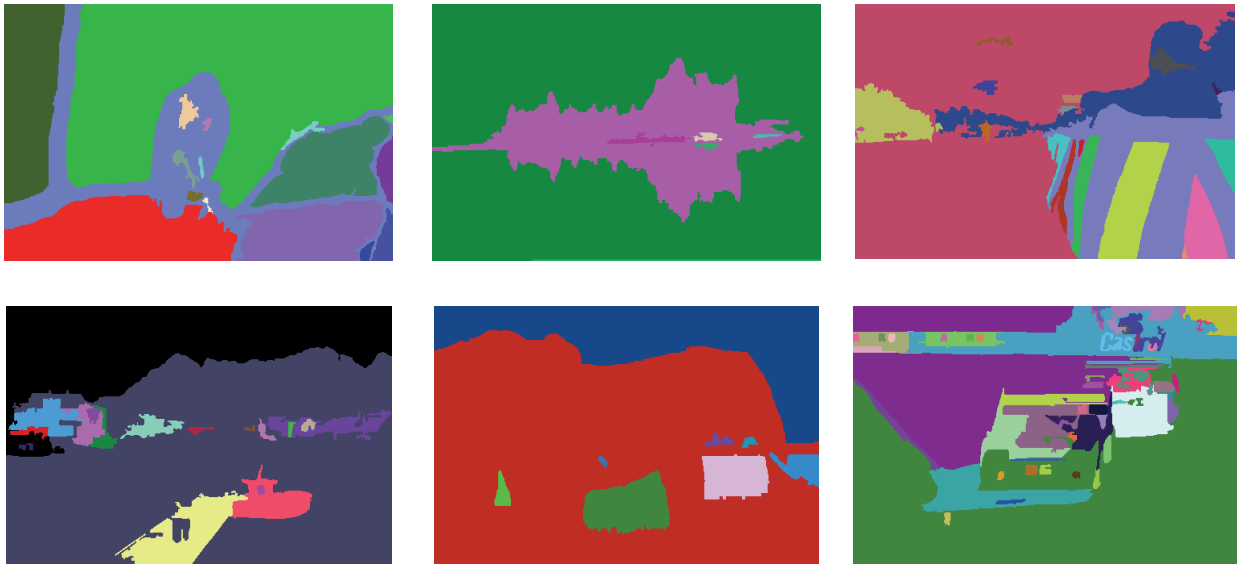




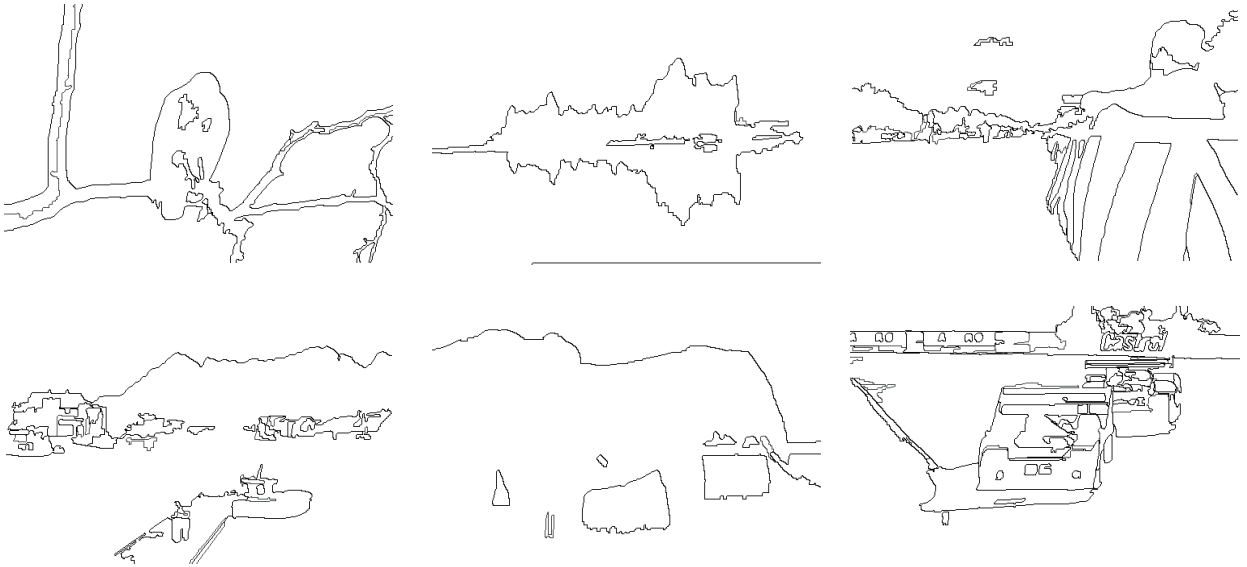
Spatial quadtree-based segmentation

- Properties
  - Low complexity: reduced number of blocks
  - Consistency via inherited labels
  - Robustness onto color gradated areas



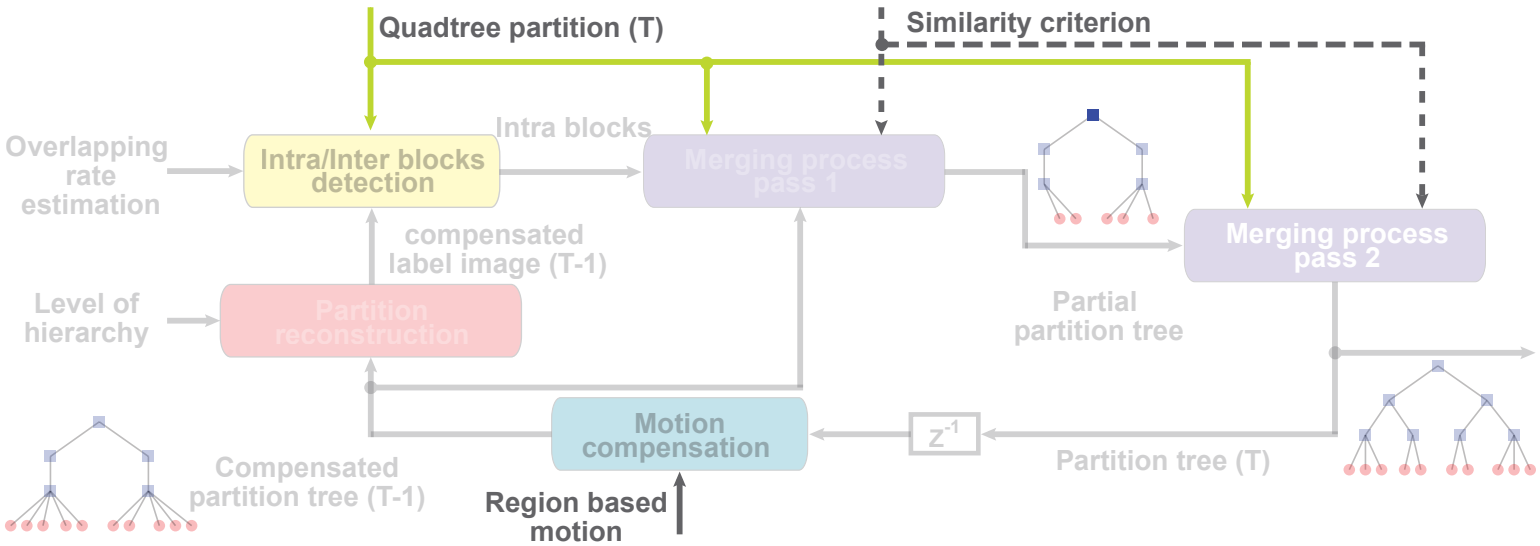
Spatial quadtree-based segmentation

- Properties
  - Low complexity: reduced number of blocks
  - Consistency via inherited labels
  - Robustness onto color gradated areas



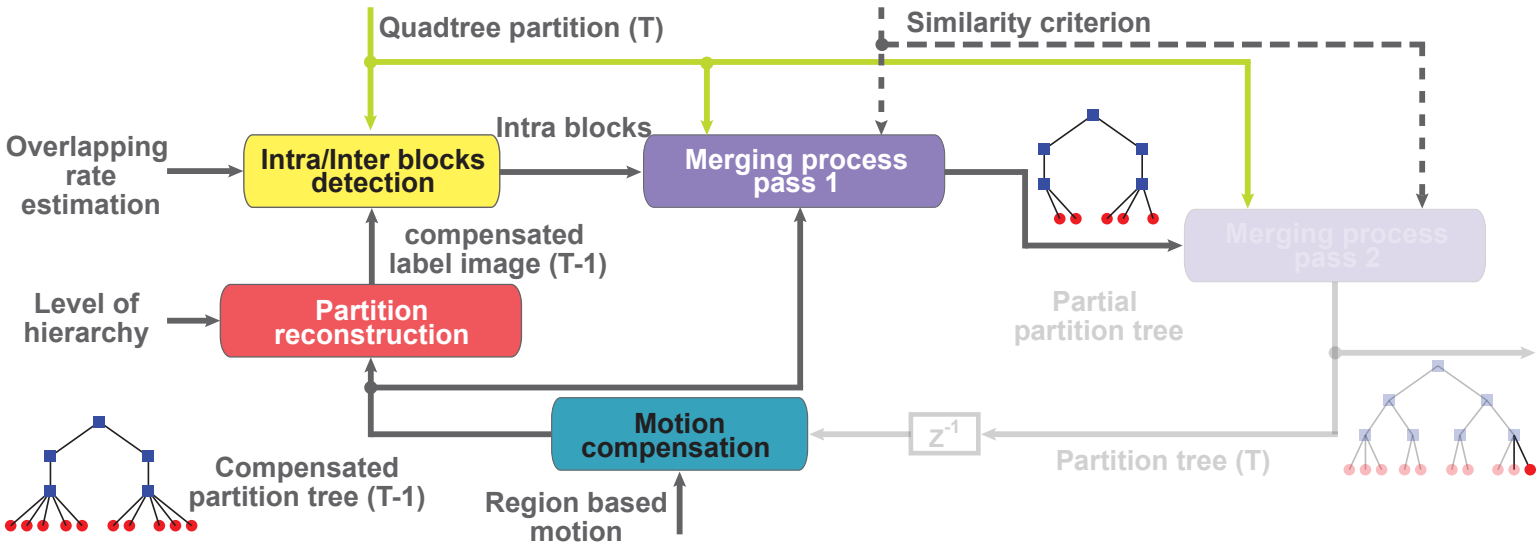
Extension to videos

- Spatio-temporal segmentation [FBD07]
  - Motion and spatial compliant hierarchy
  - Temporally stable representation
  - Two-step merging process



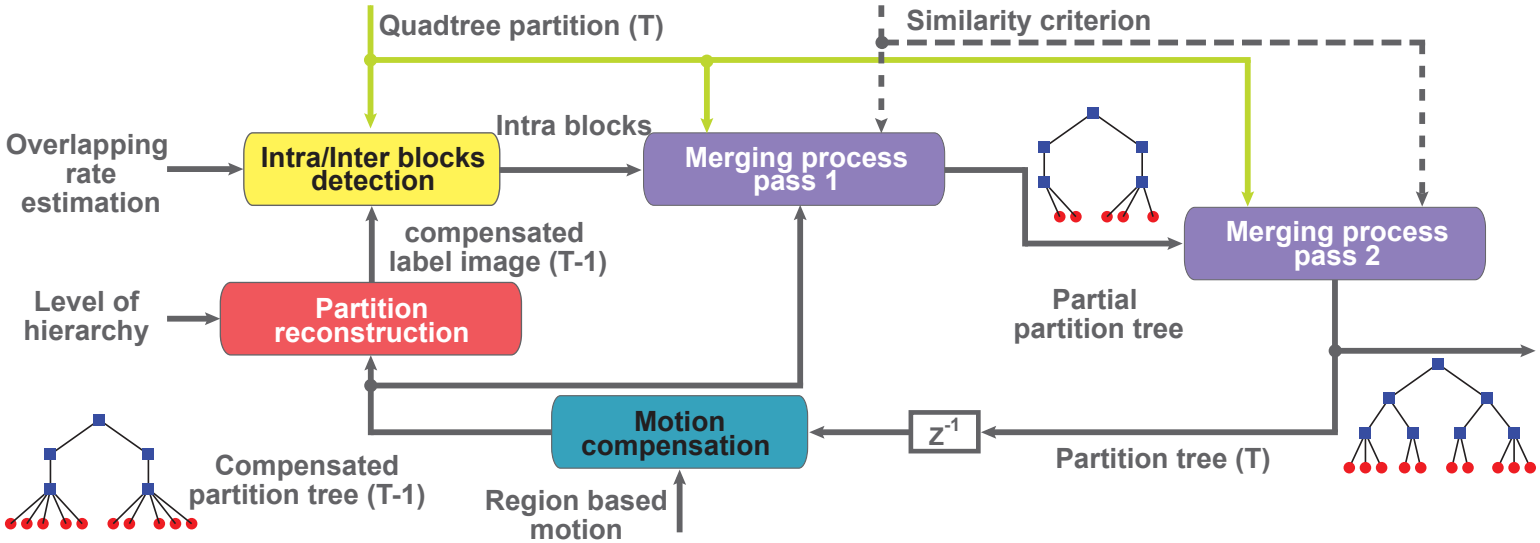
Extension to videos

- Spatio-temporal segmentation [FBD07]
  - Motion and spatial compliant hierarchy
  - Temporally stable representation
  - Two-step merging process



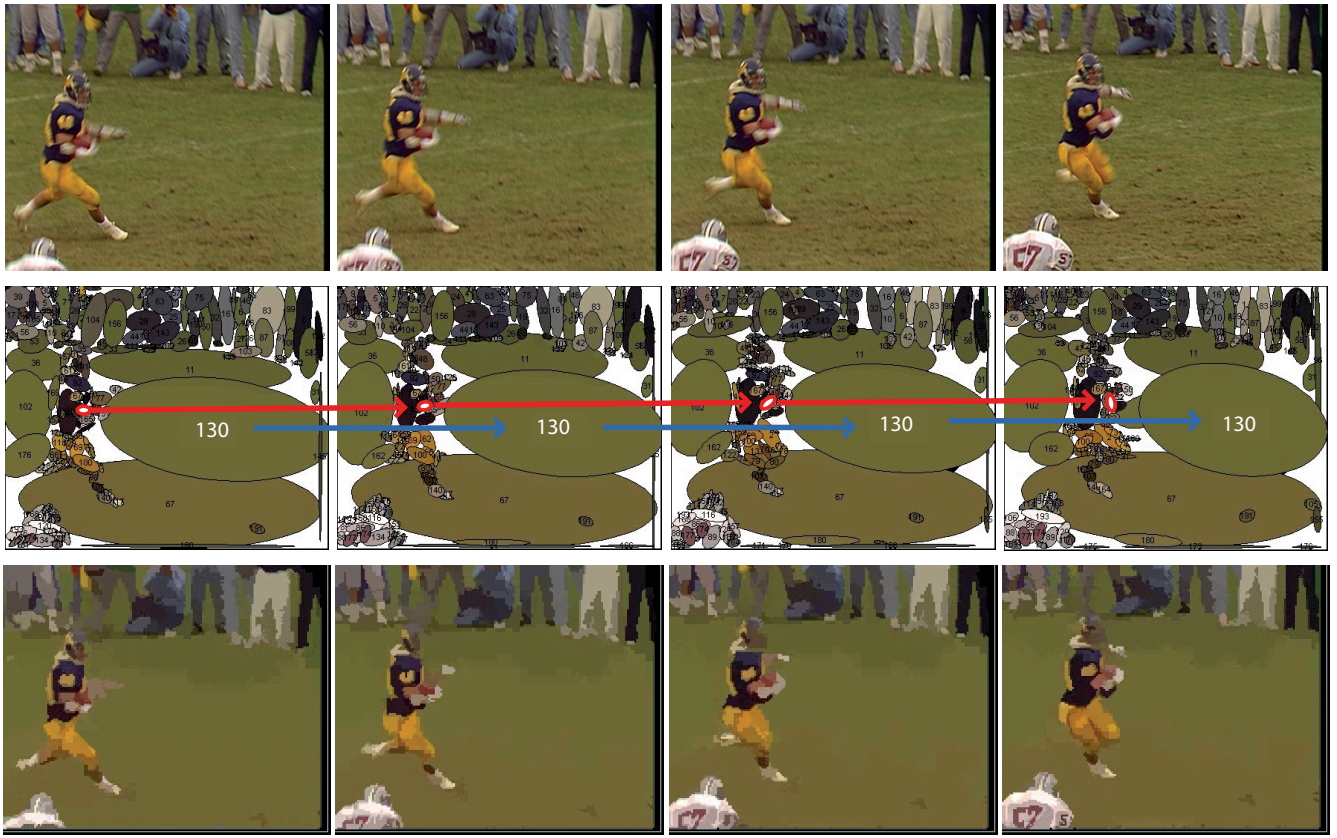
Extension to videos

- Spatio-temporal segmentation [FBD07]
  - Motion and spatial compliant hierarchy
  - Temporally stable representation
  - Two-step merging process



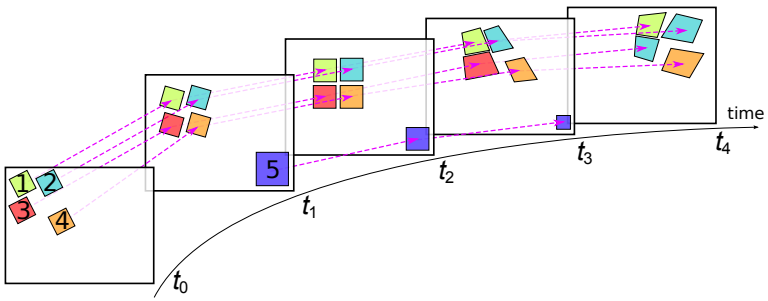
Extension to videos

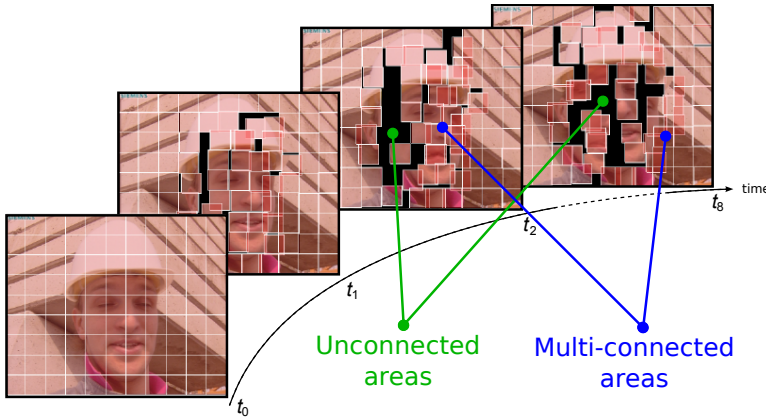
- Temporal consistency of regions



Motion tube representation

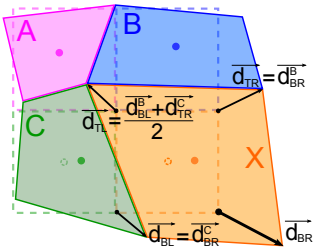
- Persistence of texture information - objects and background
- Motion tube definition [Urvoy11]
  - Moving patches of texture
  - Trajectory, deformation and lifespan models

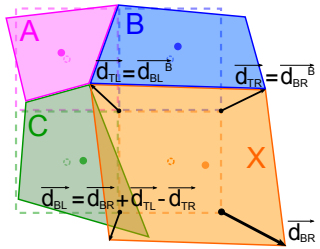


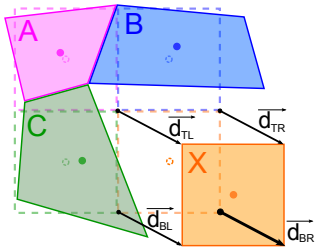


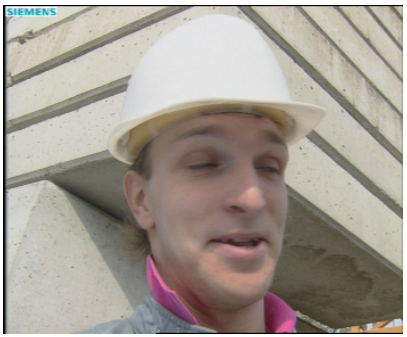
Motion tube representation

- Motion model of a tube: in between blocks and meshes
  - Modified Switched Overlapped Block Motion Compensation (SOBMC)









Original image

Disconnected blocks

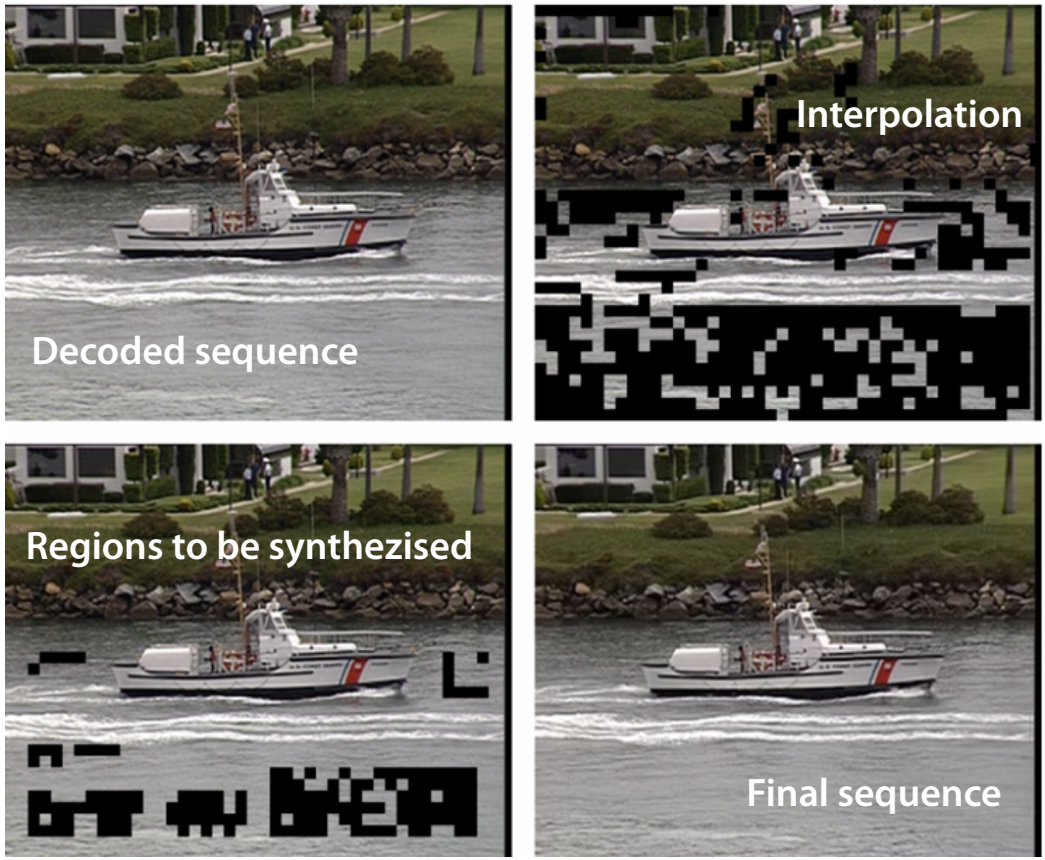
Hybrid blocks

Adaptive image synthesis

- Nearly-stationary textures
  - Exact position of texture not required for QoE purposes
  - Synthesized texture: visually relevant
- Spatio-temporal synthesis
  - Spatial context
  - Temporal consistency of textures
- Principles [Racape11]
  - Rigid regions: only motion compensation
  - Deformable textures: temporal guided synthesis
  - Switch pixel/patch based synthesis

Adaptive image synthesis

- Process overview



Pseudo-semantic representation: conclusion

- From spatial to temporal segmentation
  - Region consistency
  - Advanced semantically based processes
- Joint video analysis and coding tools
  - Disruptive technologies - Block based solutions
  - Patch or region tracking
  - Pseudo-semantic image and video representations

Pseudo-semantic representation: conclusion

- From spatial to temporal segmentation
  - Region consistency
  - Advanced semantically based processes
- Joint video analysis and coding tools
  - Disruptive technologies - Block based solutions
  - Patch or region tracking
  - Pseudo-semantic image and video representations
- Natural research topic evolution: robotic vision

Outline

Research project: towards robotic vision

- Progressive transition
  - From coding to representation systems
  - From pixel to region
- Robotic vision: related tools
  - Geometrical cue based processing
  - Object detection and texture modeling
  - Object motion and trajectory modeling
  - 3D tracking
  - Semantic scene analysis
- Particularities
  - Image processing - Segmentation
  - Embedded system considerations

Towards robotic vision

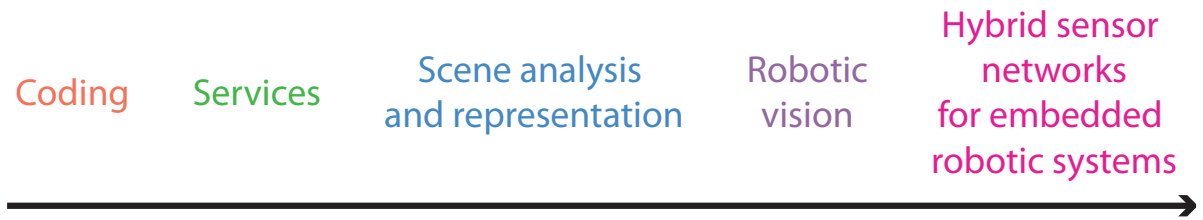
- Collaborative works
  - Personal assistance living issues: towards higher autonomy
  - Assistance au Pilotage pour l'Autonomie et la Sécurité des personnes Handicapées (APASH)
  - Pôle Images et Réseaux
  - 2012-2014
- Framework steps: motion control
  - Localization: robotic vision - robust tracking (Rafik Sekkal)
  - Navigation: data fusion and visual servoing (Zhigang Yao)
- Embedded system considerations
  - Distributed architecture
  - Ensuring system compatibility with existing electric wheelchairs

APASH project

- First step: insuring secured moves inside problematic areas
  - Going through doors, taking the elevator...
  - Avoiding collisions in dynamic indoor environment
- Desired system properties
  - No model for the environment
  - Reliability and safety
  - Man in the loop: tradeoff between autonomy and remote control
- Heterogeneous data
  - Vision sensors
  - Odometry sensors
  - Proximetry sensors

Long-term challenging issues

- Personal assistance living
  - System safety enhancement
  - Low cost and easy-to-use solutions
  - Outdoor navigation: next social and technical challenge



- Hybrid sensor systems
  - New generation of sensors: Ultra Wide Band technologies
  - Data fusion for precise localization
  - Adapting visual servoing concepts

Research activity assessment

- Supervision of 10 PhD students
- Publications
  - 4 International journals (TCSVT, CMIG, Springer JC, Springer JRTIP), 4 book chapters
  - 10 JPEG contributions, 1 patent
  - 34 International conferences (11 IEEE, 4 SPIE, 8 EUSIPCO)
- Collaborative work
  - ANR TSAR: IRCCyN, LIRMM, LIS, C2RMF (Louvre)
  - ANR CAIMAN: Thalès Communication, X-LIM, ETIS
  - APASH: AdvanSEE, Ergovie
  - Industrial collaborations: Technicolor, Orange Labs, SII, NeoTecVision
  - Academic collaborations: LIFO, HEUDIASYC, ESME, hôpital Broca
- Journal reviewer: IEEE TCSVT, IEEE TBME, Pattern Recognition, Signal Proc. Systems, IEE EL
- Conference reviewer: ICIP