



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE



Institute for
Infocomm Research

Knowledge-Based Medical Image Indexing and Retrieval

Caroline LACOSTE

Joo Hwee LIM

Jean-Pierre CHEVALLET

Daniel RACOCEANU

Nicolas Maillot

Image **P**erception, **A**ccess & **L**anguage (IPAL)
French-Singaporean Joint Lab (CNRS,I2R,NUS,UJF)



School of Computing



Approach

- **Indexing both image and text using medical concepts** from the NLM's Unified Medical Language System (UMLS)
 - To incorporate expert knowledge
 - To work at a higher semantic level
 - To unify text and image
- **Structured learning approach** to extract medical semantics from images (e.g. modality, anatomy, pathology)
- **Fusion** between textual and visual information

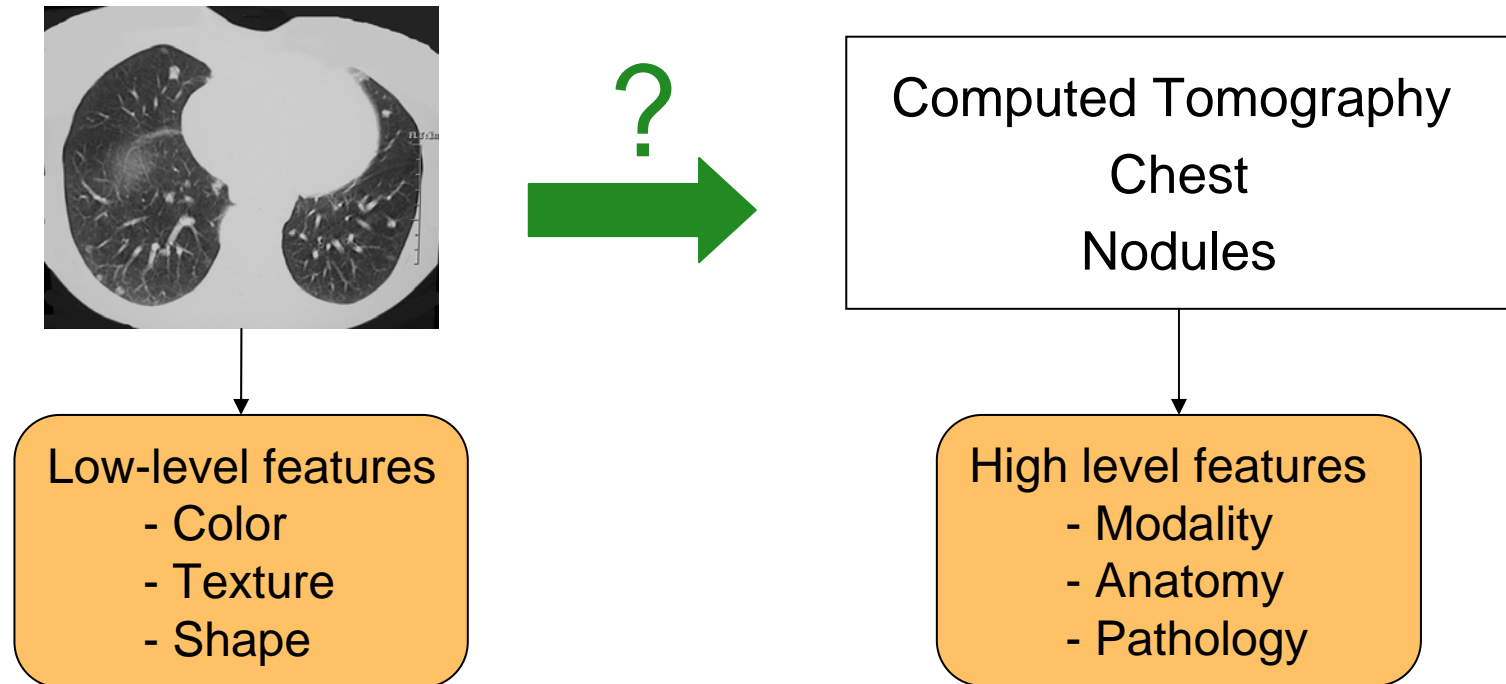


UMLS-based text indexing

- UMLS (Unified Medical Language System)
 - Multilingual: meta-thesaurus 17 languages
 - Large : More than 50000 concepts, 5,5 millions of terms
 - Consistent: categorization of concepts in semantic types
- Why using concepts and not terms ?
 - Remove the problem of term variation / synonymy
 - Ex: “Fracture” and “Broken Bones”
 - “Natural” multi-lingual indexing
- Using Meta-Thesaurus Structure
 - Vector space model does not take into account this structure
 - Using Semantic Dimension (Given by the UMLS)
- Tested this year
 - **Dimension filtering**: At least one matching according one of the three dimension
 - **Dimension weighting**: Re-weight similarity according to the number of matched dimension



Semantic Image Indexing

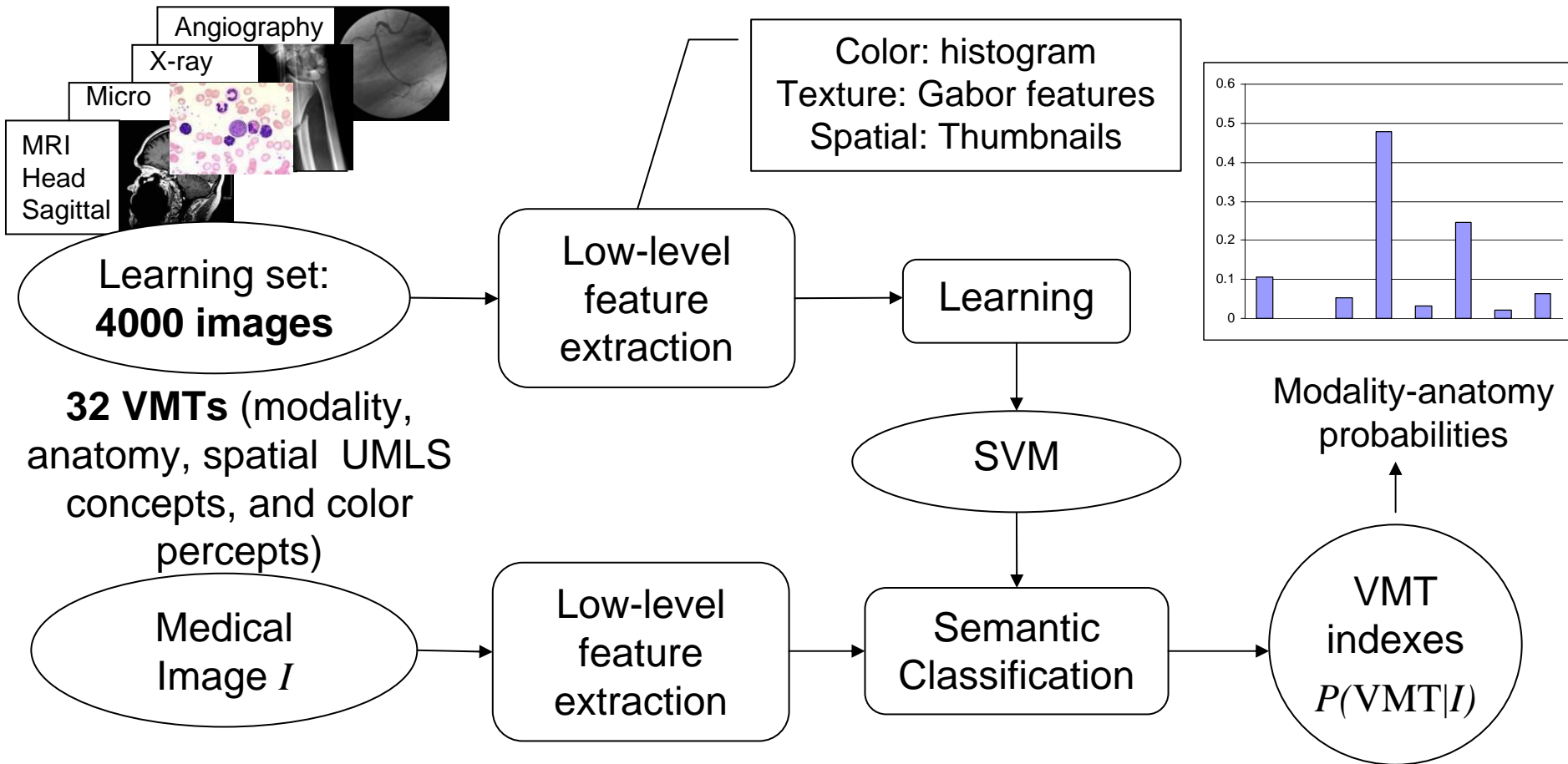


Supervised learning framework

- Global indexing to access image modality
- Local indexing to access semantic local features related to anatomy, modality, and pathology concepts



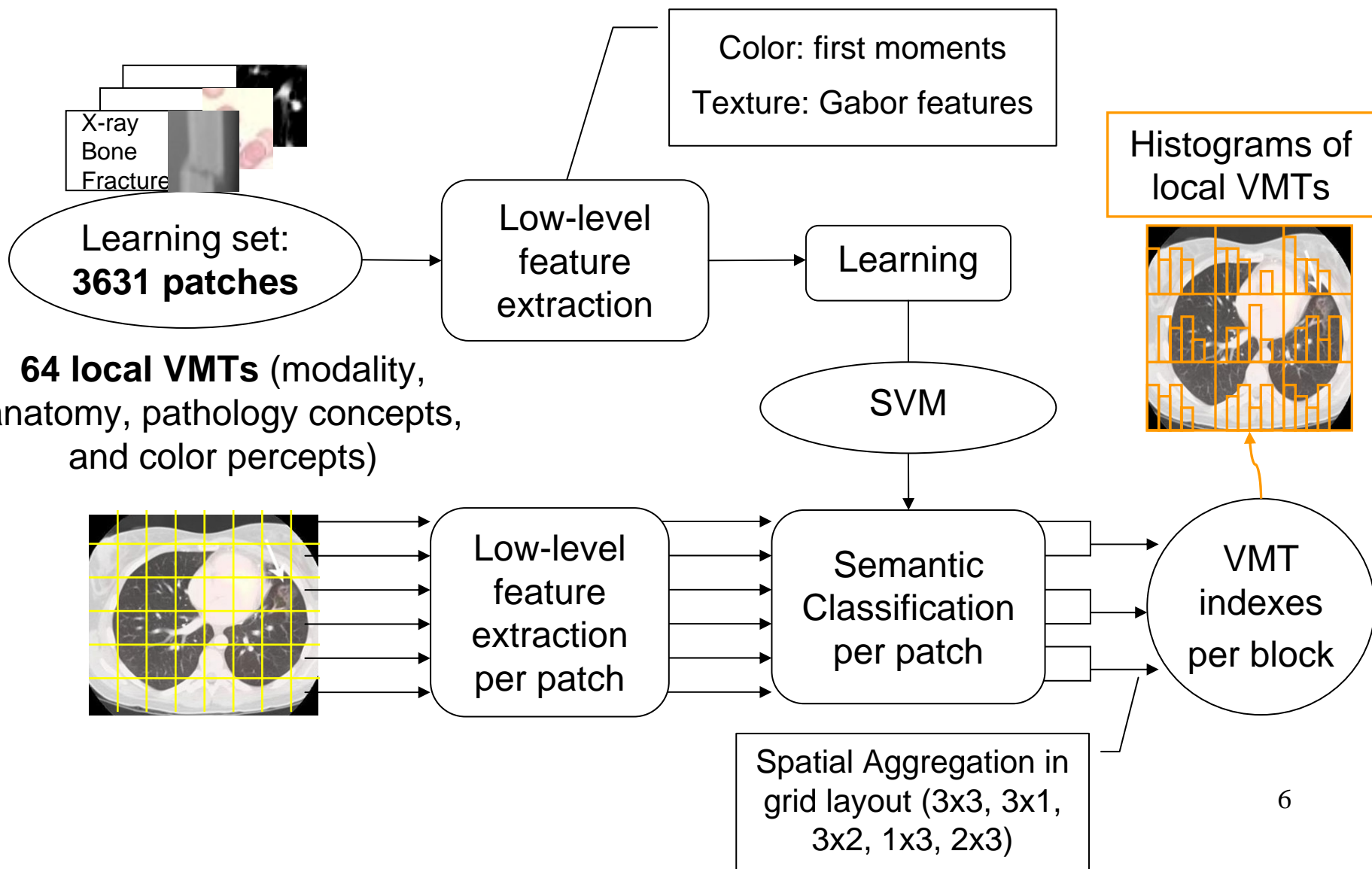
Global Indexing



VMT: Visual Medical Terms



Local indexing





Visual retrieval

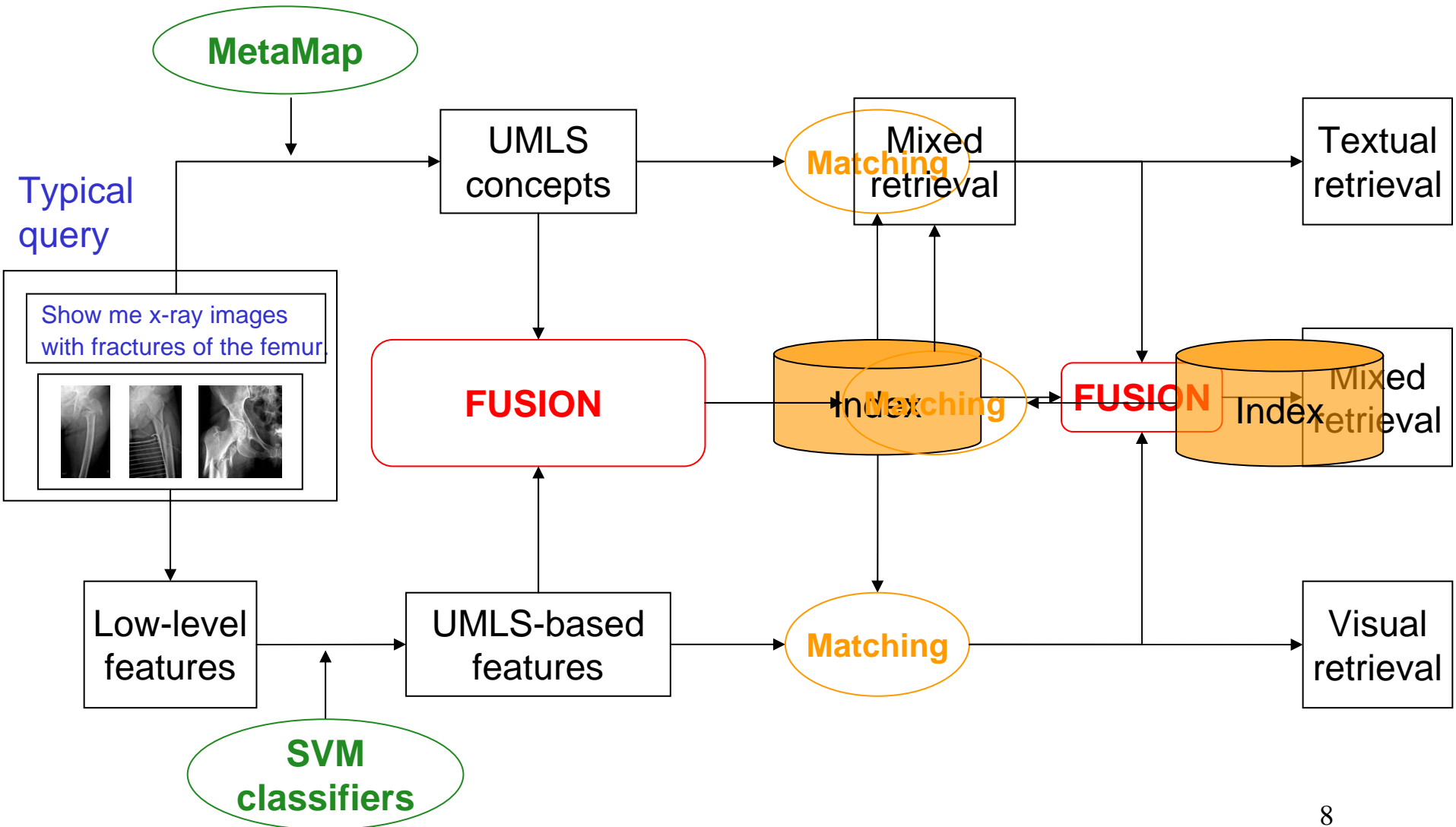
- Retrieval using the global visual indexing
 - Retrieval based on the Manhattan distance between 2 indexes (~histogram of modality-anatomy concepts)
 - Modality filtering according to the textual query modality concepts
 - I admissible if $P(mod_Q / I) > T$

- Retrieval using the local indexing
 - Retrieval based on the mean of Manhattan distances between local VMT histograms

- Late Fusion : mean of the two similarity measures



Medical Image Retrieval





Comparative Results on CLEF2005

Method	Textual	Visual	MAP	Rank
Concept, Dimension re-weighting (1)	x		26.46%	1/31
Concept, Pseudo-Rel. FB, Dim filt	x		22.94%	2/31
Concepts, Dimension filtering	x		22.70%	3/31
Terms, Dimension filtering	x		20.88%	5/31
Concept, semantic doc expansion	x		18.56%	10/31
Fusion Local+Global Indexing (2)		x	06.41%	3/11
Global Indexing according to modality		x	05.66%	4/11
Local Indexing with VMT		x	04.84%	6/11
Late fusion (1) + (2)	x	x	30.95%	1/37
Concept fusion (1) + (2) adaptive threshold	x	x	28.78%	2/37
Concept fusion (1) + (2) Fixed Threshold 0.15	x	x	28.45%	3/37



Conclusion

- CLEF multilingual medical image retrieval task:
 - Very difficult task: large collection, “precise” and “semantic” queries
- Concept Indexing:
 - It works !!
 - Inter-media common indexing
 - Multilingual
- Image and text: complementary
 - Text: closer to the meaning
 - Efficiency of the visual filtering to remove aberrant images
- Importance of
 - semantic dimensions for text
 - visual terms and learning
- The UMLS indexing induces a lot of perspectives:
 - Early fusion
 - Semantic-based retrieval



Inter-Media Pseudo-Relevance Feedback Application to ImageCLEF Photo 2006

Nicolas MAILLOT
Jean-Pierre CHEVALLET
Joo Hwee LIM

Image **P**erception, **A**ccess & **L**anguage (IPAL)
French-Singaporean Joint Lab (CNRS,I2R,NUS,UJF)



Inter-Media Pseudo-Relevance Feedback

- Problem:
 - Multi-modal (text+image) Information Indexing and Retrieval
 - Multi-modal documents
 - Multi-modal queries
 - Application to the ImageCLEF Photo Task

- Objectives:
 - **Inter-media enrichment**
 - ➡ dealing with synonymy
 - **Re-using** existing mono-media image and text retrieval systems



Inter-Media Pseudo-Relevance Feedback

Related Works

- Mono-modal pseudo-relevance feedback [*Xu and Croft 96*]
- Translation models [*Lin et al. 05*]
 - Automatic translation of the textual query into a visual query
 - Requires prior mining of textual/visual relationships
- Late Fusion [*Chevallet et al. 05*]
 - Text and image retrieval are done in parallel and the results are merged (weighted sum)
 - No mutual enrichment
- Latent Semantic Indexing (text+visual keywords)
 - Computationally expensive



Inter-Media Pseudo-Relevance Feedback

Text Processing

- Morpho-Syntax
 - Part of Speech Tagging
 - Unknown Proper Nouns detection
 - Word Normalization & Spelling correction
- Index terms
 - Noun Phrase detection using WordNet
 - Geographic named entities detected using Wordnet or the <LOCATION> tag
- Concept Indexing
 - Selection based on the most frequent sense provided by Wordnet



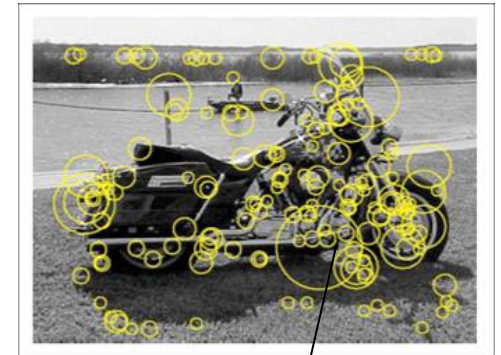
Inter-Media Pseudo-Relevance Feedback

Image Processing

- Image Segmentation
 - Meanshift Segmentation
 - Patch Based Tessellation



- Feature Extraction
 - Color histograms
 - Bags of SIFT
 - Scale Invariant Feature Transform
 - Gabor Features (Texture)

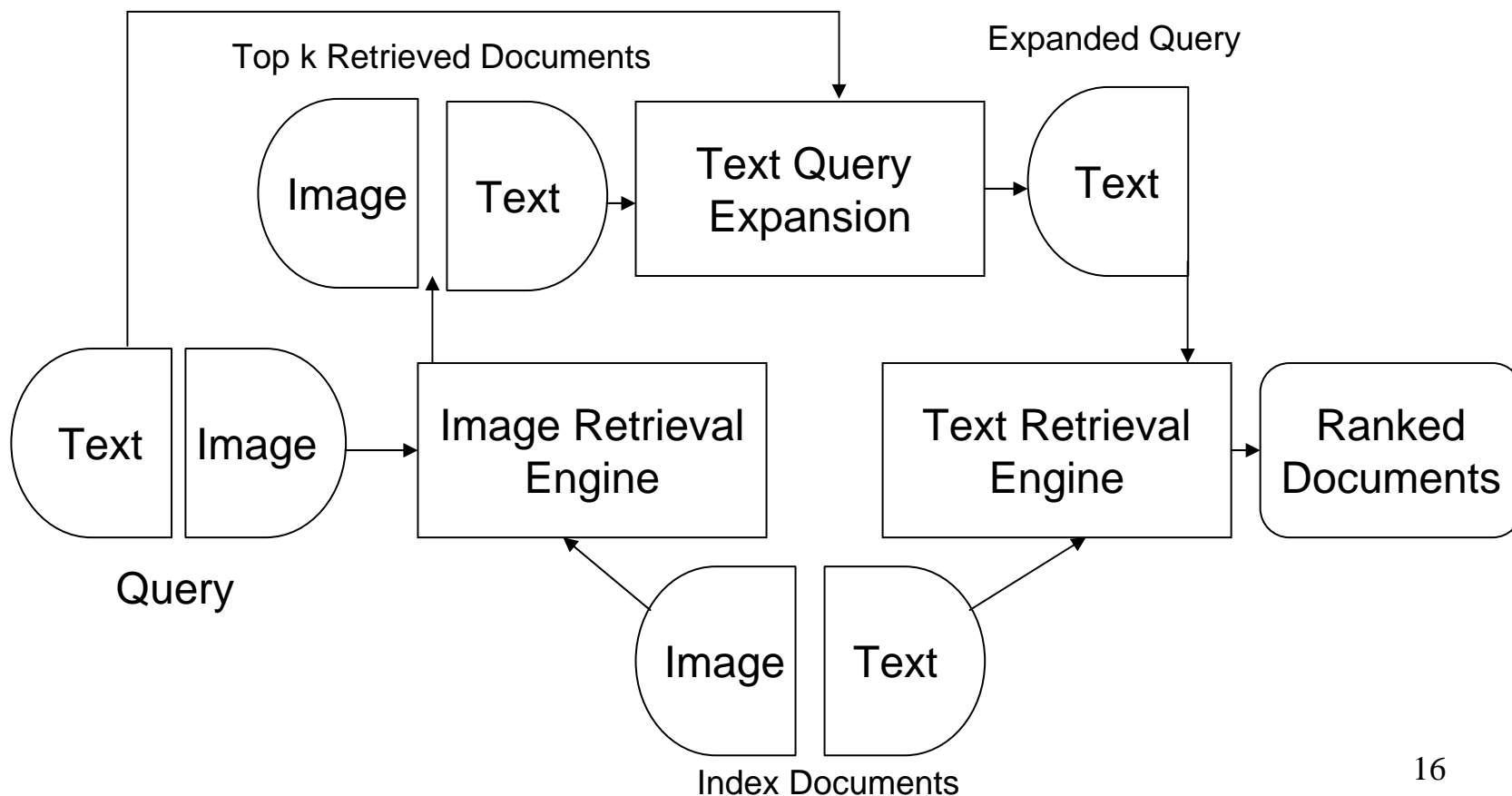


Extraction of one local
orientation histogram per
location



Inter-Media Pseudo-Relevance Feedback

Overview





Inter-Media Pseudo-Relevance Feedback

Results

	Modality	A/M	MAP	P10	P20	P30	Rank
Inter Media Pseudo Relevance Feedback	Text+Visual	Automatic	0.3337	0.5067	0.5000	0.4583	2/43
Best Run	Text+Visual	Manual	0.3850	0.6283	0.5300	0.4550	1/43



Inter-Media Pseudo-Relevance Feedback

Conclusion

- Using the image modality to expand the text query
 - Use of the text associated with the top k images retrieved
 - Assumption: these k images are relevant
- Future work:
 - Advanced conceptual filtering and reasoning
 - some concepts are not characterized by the visual appearance
 - Comparison with the translation model



Thanks !



Supervised learning framework

