Video Object Mining : Issues and Perspectives

Jonathan Weber, Sébastien Lefèvre, Pierre Gançarski

LSIIT,University of Strasbourg
Ready Business System

September 23, 2010

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

$\Longrightarrow$ Lack of semantic information

Issues

High quantity of video data

- Low cost digital camera
- High storage capacity
- High-speed internet
- Ex : YouTube, Dailymotion, ...

How can we mine efficiently these complex data ?

User needs content-based tools but

- Video meta-data are not sufficient
- Manual tagging of video is time consuming and error-prone
- There is a semantic gap between video data and what they represent

$\Longrightarrow$ Lack of semantic information

How can we introduce semantic in the mining process ?

Video Mining Systems

Video Mining

- Process of extracting information/knowledge from large amounts of video data
- Many recent publications deal with video mining:
  - A. Anjulan and N. Canagarajah, Signal Processing: Image Communication, 2007
  - A. Anjulan and N. Canagarajah, IEEE ICIP, 2007
  - S. de Avila, A. da Luz, and A. de Araujo, IWSSIP, 2008
  - A. Basharat, Y. Zhai, and M. Shah, Computer Vision and Image Understanding, 2008
  - F. Chevalier, J.-P. Domenger, J. Benois-Pineau, and M. Delest, Pattern Recognition Letters, 2007
  - X. Gao, X. Li, J. Feng, and D. Tao, Pattern Recognition Letters, 2009
  - D. Liu and T. Chen, Computer Vision and Image Understanding, 2009
  - W. Ren and Y. Zhu, IIH-MSP, 2008
  - J. Sivic and A. Zisserman, Proceedings of the IEEE, 2008
  - L. F. Teixeira and L. Corte-Real, Pattern Recognition Letters, 2009
  - S. Zhai, B. Luo, J. Tang, and C.-Y. Zhang, ICIG, 2007

Study of existing systems

- Characterize the different systems
- Check if they use semantics

Video Mining Systems

Video Mining

- Process of extracting information/knowledge from large amounts of video data
- Many recent publications deal with video mining:
  - A. Anjulan and N. Canagarajah, Signal Processing: Image Communication, 2007
  - A. Anjulan and N. Canagarajah, IEEE ICIP, 2007
  - S. de Avila, A. da Luz, and A. de Araujo, IWSSIP, 2008
  - A. Basharat, Y. Zhai, and M. Shah, Computer Vision and Image Understanding, 2008
  - F. Chevalier, J.-P. Domenger, J. Benois-Pineau, and M. Delest, Pattern Recognition Letters, 2007
  - X. Gao, X. Li, J. Feng, and D. Tao, Pattern Recognition Letters, 2009
  - D. Liu and T. Chen, Computer Vision and Image Understanding, 2009
  - W. Ren and Y. Zhu, IIH-MSP, 2008
  - J. Sivic and A. Zisserman, Proceedings of the IEEE, 2008
  - L. F. Teixeira and L. Corte-Real, Pattern Recognition Letters, 2009
  - S. Zhai, B. Luo, J. Tang, and C.-Y. Zhang, ICIG, 2007

Study of existing systems

- Characterize the different systems
- Check if they use semantics

Video Mining Systems

Video Mining

- Process of extracting information/knowledge from large amounts of video data
- Many recent publications deal with video mining:
  - A. Anjulan and N. Canagarajah, Signal Processing: Image Communication, 2007
  - A. Anjulan and N. Canagarajah, IEEE ICIP, 2007
  - S. de Avila, A. da Luz, and A. de Araujo, IWSSIP, 2008
  - A. Basharat, Y. Zhai, and M. Shah, Computer Vision and Image Understanding, 2008
  - F. Chevalier, J.-P. Domenger, J. Benois-Pineau, and M. Delest, Pattern Recognition Letters, 2007
  - X. Gao, X. Li, J. Feng, and D. Tao, Pattern Recognition Letters, 2009
  - D. Liu and T. Chen, Computer Vision and Image Understanding, 2009
  - W. Ren and Y. Zhu, IIH-MSP, 2008
  - J. Sivic and A. Zisserman, Proceedings of the IEEE, 2008
  - L. F. Teixeira and L. Corte-Real, Pattern Recognition Letters, 2009
  - S. Zhai, B. Luo, J. Tang, and C.-Y. Zhang, ICIG, 2007

Study of existing systems

- Characterize the different systems
- Check if they use semantics

Video Mining Systems

Video Mining

- Process of extracting information/knowledge from large amounts of video data
- Many recent publications deal with video mining:
  - A. Anjulan and N. Canagarajah, Signal Processing: Image Communication, 2007
  - A. Anjulan and N. Canagarajah, IEEE ICIP, 2007
  - S. de Avila, A. da Luz, and A. de Araujo, IWSSIP, 2008
  - A. Basharat, Y. Zhai, and M. Shah, Computer Vision and Image Understanding, 2008
  - F. Chevalier, J.-P. Domenger, J. Benois-Pineau, and M. Delest, Pattern Recognition Letters, 2007
  - X. Gao, X. Li, J. Feng, and D. Tao, Pattern Recognition Letters, 2009
  - D. Liu and T. Chen, Computer Vision and Image Understanding, 2009
  - W. Ren and Y. Zhu, IIH-MSP, 2008
  - J. Sivic and A. Zisserman, Proceedings of the IEEE, 2008
  - L. F. Teixeira and L. Corte-Real, Pattern Recognition Letters, 2009
  - S. Zhai, B. Luo, J. Tang, and C.-Y. Zhang, ICIG, 2007

Study of existing systems

- Characterize the different systems
- Check if they use semantics

Essential properties to characterize video mining systems

Properties

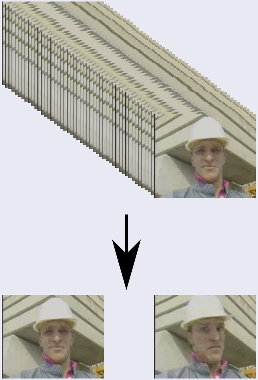
- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Summarization

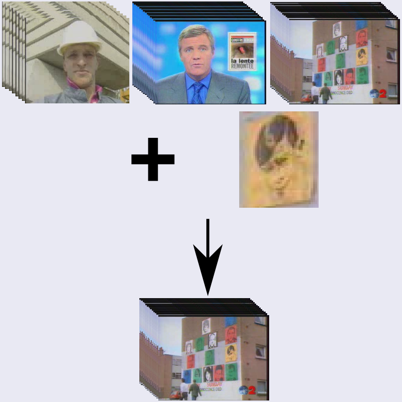


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Retrieval

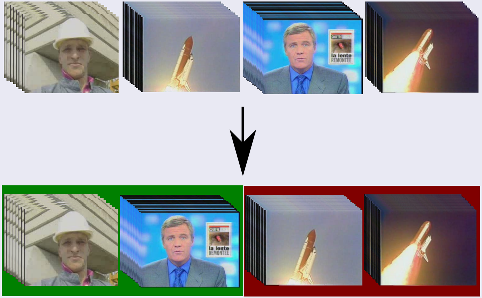


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Classification



Essential properties to characterize video mining systems

Properties

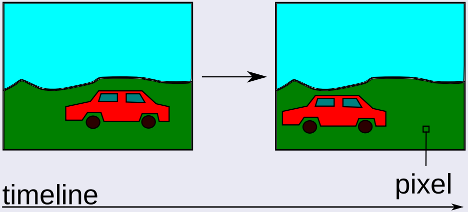
- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Pixel

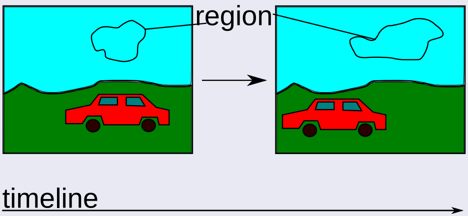


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Region

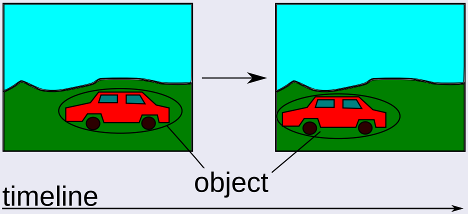


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Object

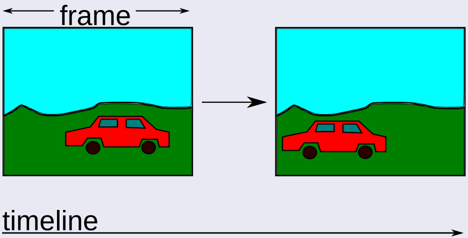


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Frame

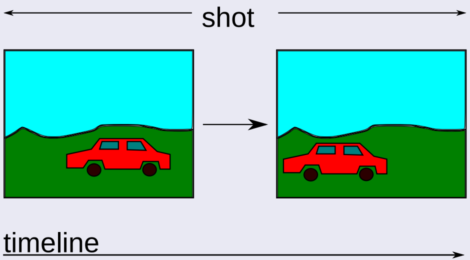


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Shot



Essential properties to characterize video mining systems

Properties

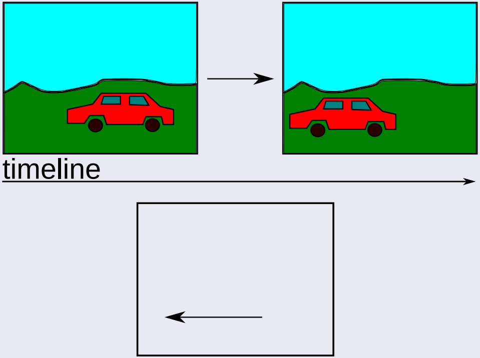
- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Motion

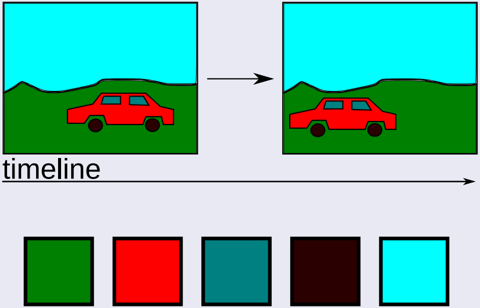


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Color

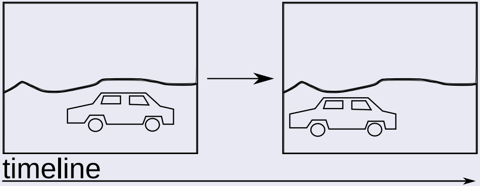


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Shape



Essential properties to characterize video mining systems

Properties

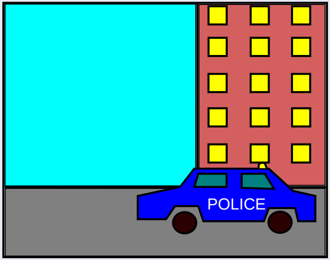
- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Global

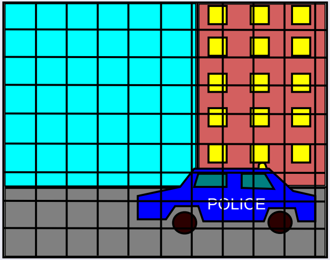


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Grid

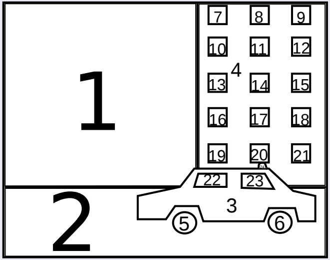


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Region

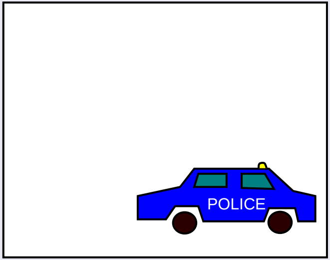


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Object

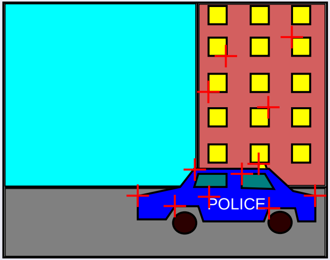


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Interest point

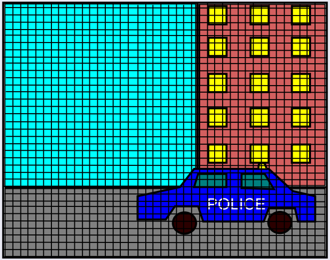


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Pixel



Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

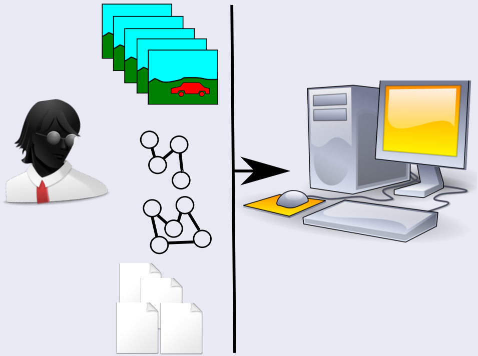
Null

Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Supervised

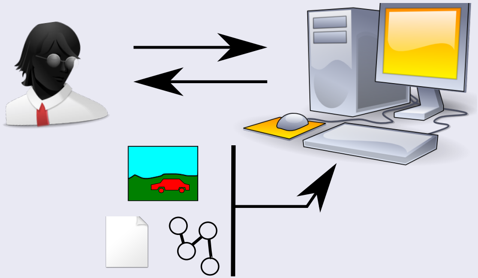


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Semi-Supervised

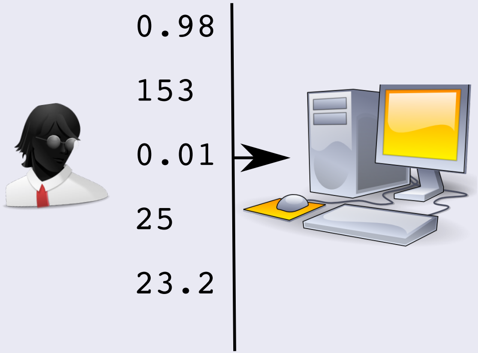


Essential properties to characterize video mining systems

Properties

- Objectives
- Elements
- Descriptors
- Scales
- Supervision levels

Parameters



Current usages of 11 existing systems (2007-2009)

Objectives

Elements

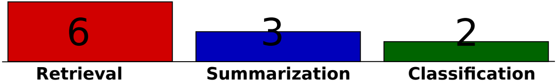
Descriptors

Scales

Supervision

Current usages of 11 existing systems (2007-2009)

Objectives



Elements

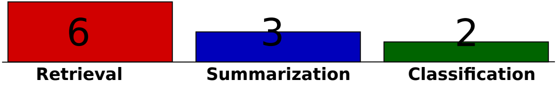
Descriptors

Scales

Supervision

Current usages of 11 existing systems (2007-2009)

Objectives



Elements

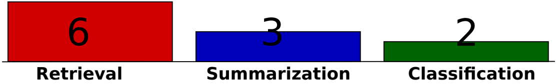
Descriptors

Scales

Supervision

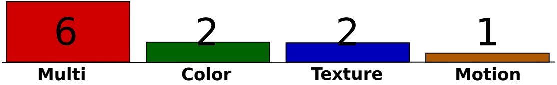
Current usages of 11 existing systems (2007-2009)

Objectives



Elements

Descriptors

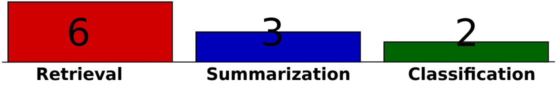


Scales

Supervision

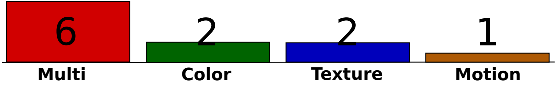
Current usages of 11 existing systems (2007-2009)

Objectives

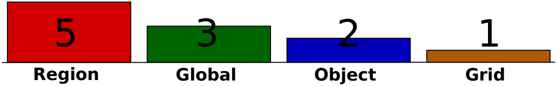


Elements

Descriptors

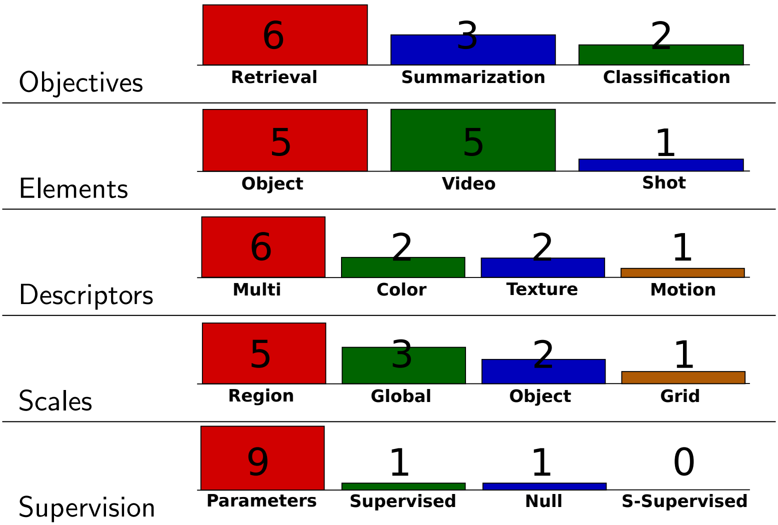


Scales



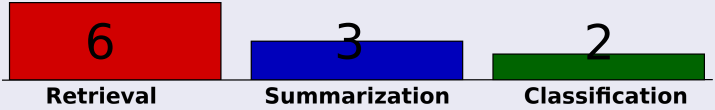
Supervision

Current usages of 11 existing systems (2007-2009)



Video Object Mining: A way of introducing semantics

Objectives



Our approach for Video Object Mining

- Generic framework
- Object is the most semantic element
- Combination of object-oriented descriptors
- 2 types of scales : inner and context scale
- User relevance feedback brings semantics

Video Object Mining: A way of introducing semantics

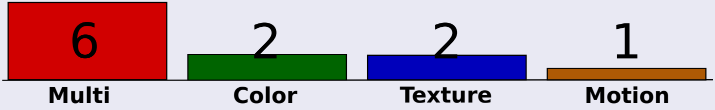
Elements

Our approach for Video Object Mining

- Generic framework
- Object is the most semantic element
- Combination of object-oriented descriptors
- 2 types of scales : inner and context scale
- User relevance feedback brings semantics

Video Object Mining: A way of introducing semantics

Descriptors

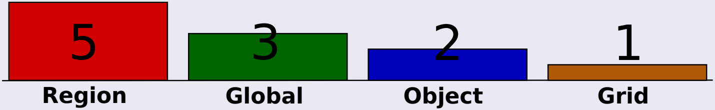


Our approach for Video Object Mining

- Generic framework
- Object is the most semantic element
- Combination of object-oriented descriptors
- 2 types of scales : inner and context scale
- User relevance feedback brings semantics

Video Object Mining: A way of introducing semantics

Scales

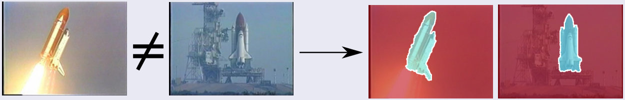


Our approach for Video Object Mining

- Generic framework
- Object is the most semantic element
- Combination of object-oriented descriptors
- 2 types of scales : inner and context scale
- User relevance feedback brings semantics

Video Object Mining: A way of introducing semantics

Scales



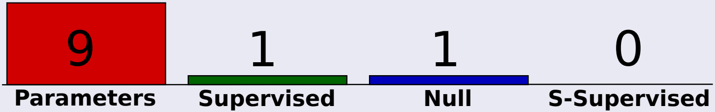
inner
context

Our approach for Video Object Mining

- Generic framework
- Object is the most semantic element
- Combination of object-oriented descriptors
- 2 types of scales : inner and context scale
- User relevance feedback brings semantics

Video Object Mining: A way of introducing semantics

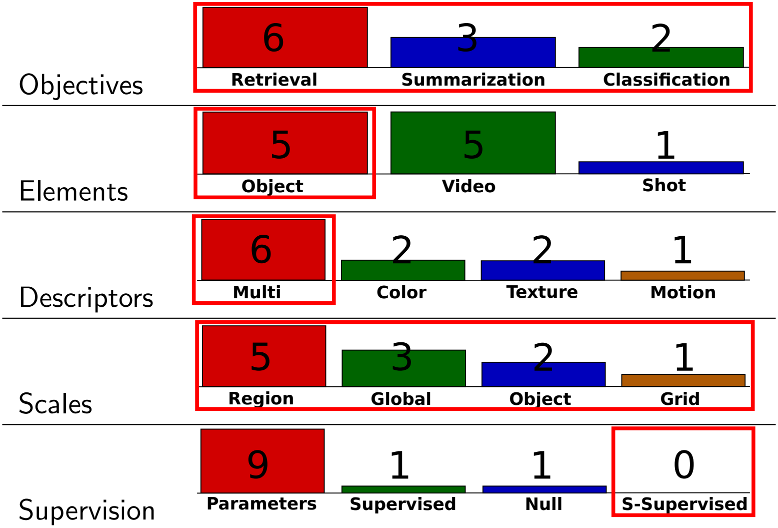
Supervision level



Our approach for Video Object Mining

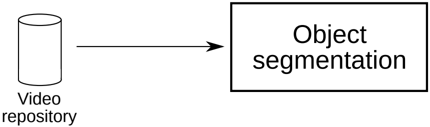
- Generic framework
- Object is the most semantic element
- Combination of object-oriented descriptors
- 2 types of scales : inner and context scale
- User relevance feedback brings semantics

Our approach compared to related works

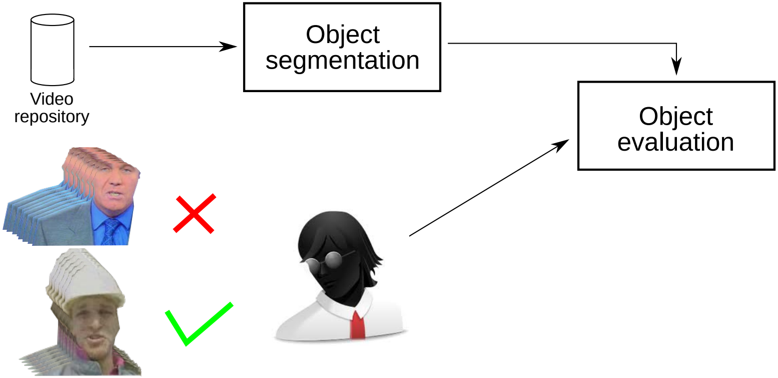


VOMF: Video Object Mining Framework

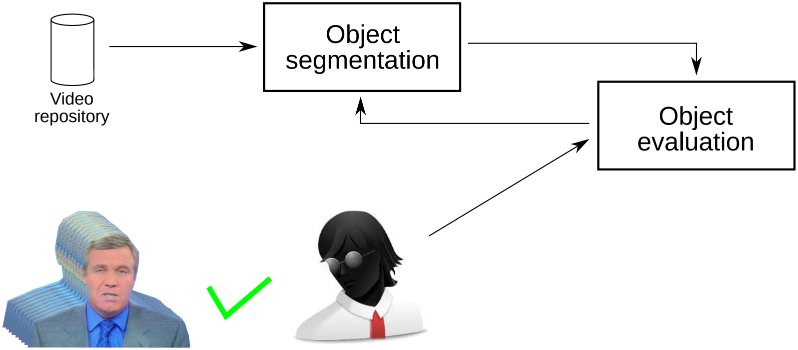
VOMF: Video Object Mining Framework



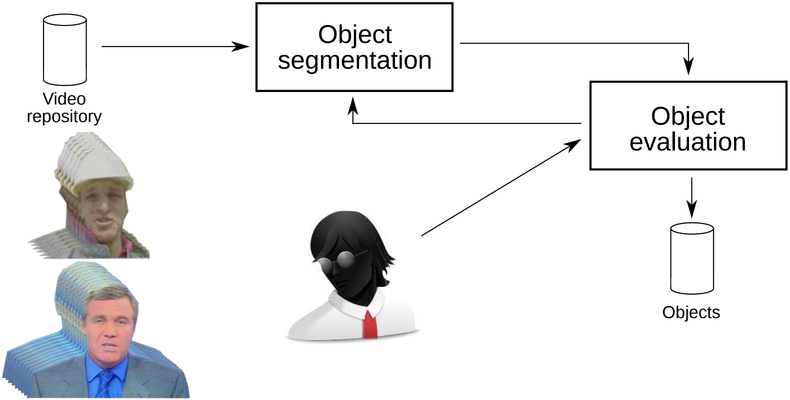
VOMF: Video Object Mining Framework



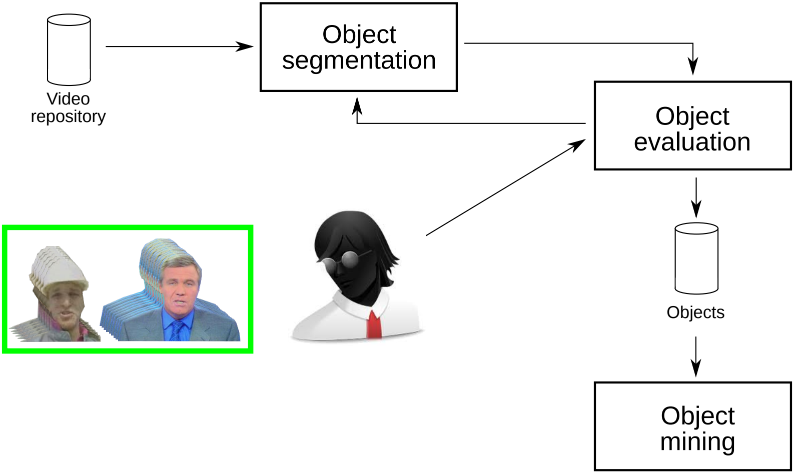
VOMF: Video Object Mining Framework



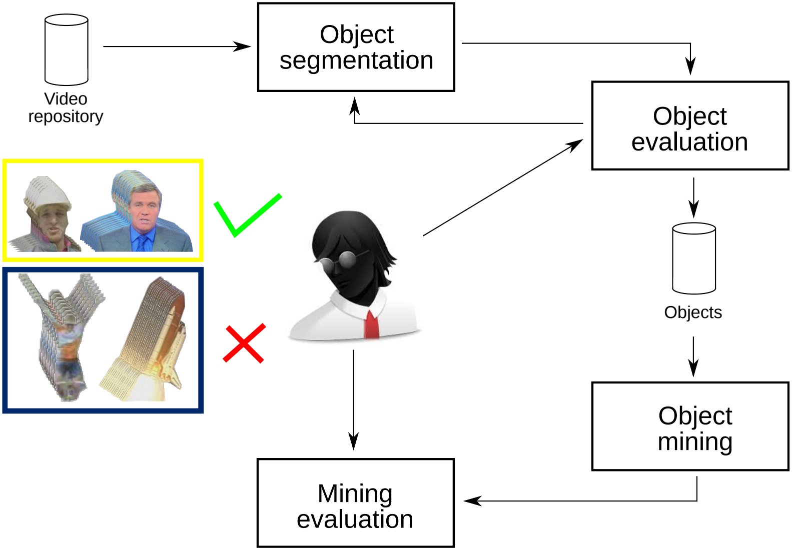
VOMF: Video Object Mining Framework



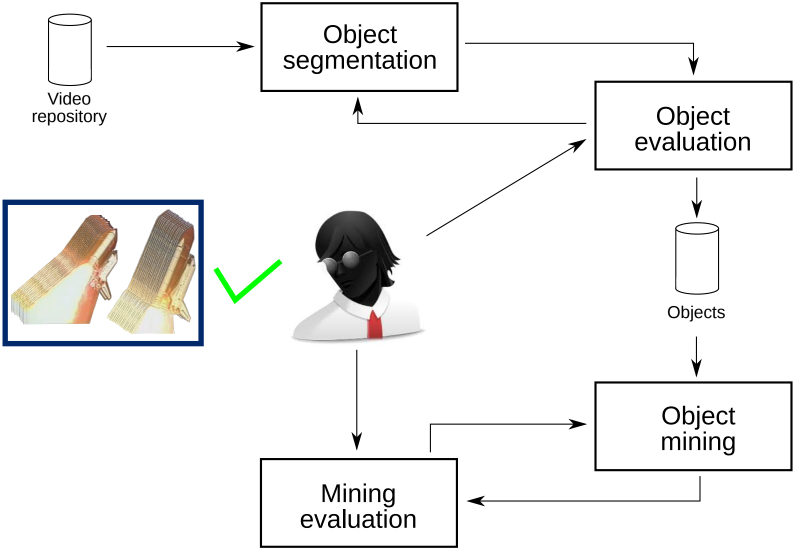
VOMF: Video Object Mining Framework



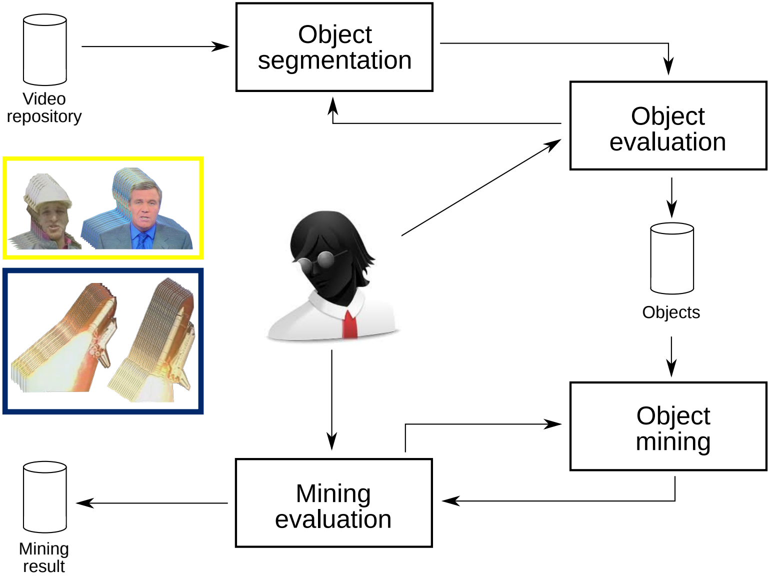
VOMF: Video Object Mining Framework



VOMF: Video Object Mining Framework



VOMF: Video Object Mining Framework



Conclusion

Issues

- Massive video repositories need new data mining schemes
- Semantics have to be considered to achieve user goals
- These are still open problems

Perspectives

- Video Object Mining has to be explored
- Semi-supervised learning limits user's workload
- A new generic framework has been proposed
- *This framework is currently being implemented*

Difficulties

- Object segmentation
- Interactive mining (involving user relevance feedback)

Conclusion

Issues

- Massive video repositories need new data mining schemes
- Semantics have to be considered to achieve user goals
- These are still open problems

Perspectives

- Video Object Mining has to be explored
- Semi-supervised learning limits user's workload
- A new generic framework has been proposed
- *This framework is currently being implemented*

Difficulties

- Object segmentation
- Interactive mining (involving user relevance feedback)

Conclusion

Issues

- Massive video repositories need new data mining schemes
- Semantics have to be considered to achieve user goals
- These are still open problems

Perspectives

- Video Object Mining has to be explored
- Semi-supervised learning limits user's workload
- A new generic framework has been proposed
- *This framework is currently being implemented*

Difficulties

- Object segmentation
- Interactive mining (involving user relevance feedback)

Conclusion

Issues

- Massive video repositories need new data mining schemes
- Semantics have to be considered to achieve user goals
- These are still open problems

Perspectives

- Video Object Mining has to be explored
- Semi-supervised learning limits user's workload
- A new generic framework has been proposed
- *This framework is currently being implemented*

Difficulties

- Object segmentation
- Interactive mining (involving user relevance feedback)