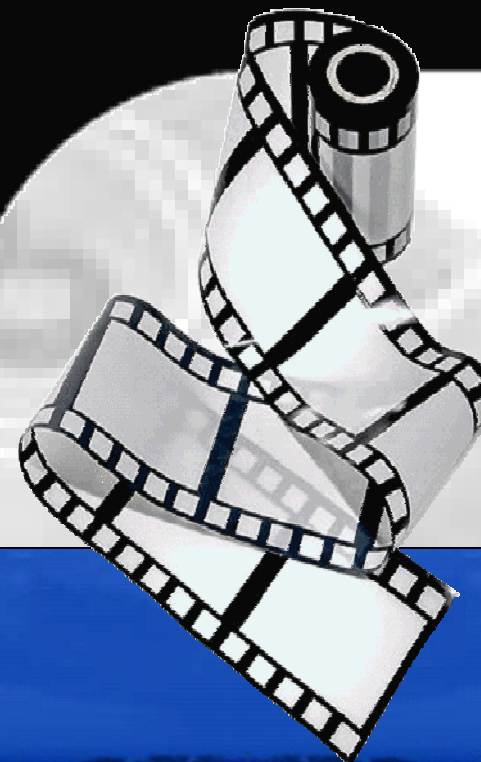
June 26, 2006

A Multi-Level Approach for Temporal Video Segmentation based on Adaptive Examples

© Robert Babak Yeganeh

Short Version



Submitted to the Department of Electrical Engineering and Computer Science and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science in Computer Science.

Committee

Dr. John Gauch (Chair)

Dr. Arvin Agah

Dr. James Miller

June 26, 2006

A Multi-Level Approach for Video Temporal Segmentation based on Adaptive Examples

Outline

1. Overview
2. Literature Review
3. Segmentation based on Predefined Examples (S.P.E.)
4. Segmentation based on Adaptive Examples (S.A.E.)
5. Experimentation Results
6. Conclusions

1. Overview

1. Overview

1. Overview

Outline

1. Overview
2. Literature Review
3. Segmentation based on Predefined Examples (S.P.E.)
4. Segmentation based on Adaptive Examples (S.A.E.)
5. Experimentation Results
6. Conclusions

2. Literature Review

- Representation
- Detection
- Classification
- False Detection and Prevention

2. Literature Review

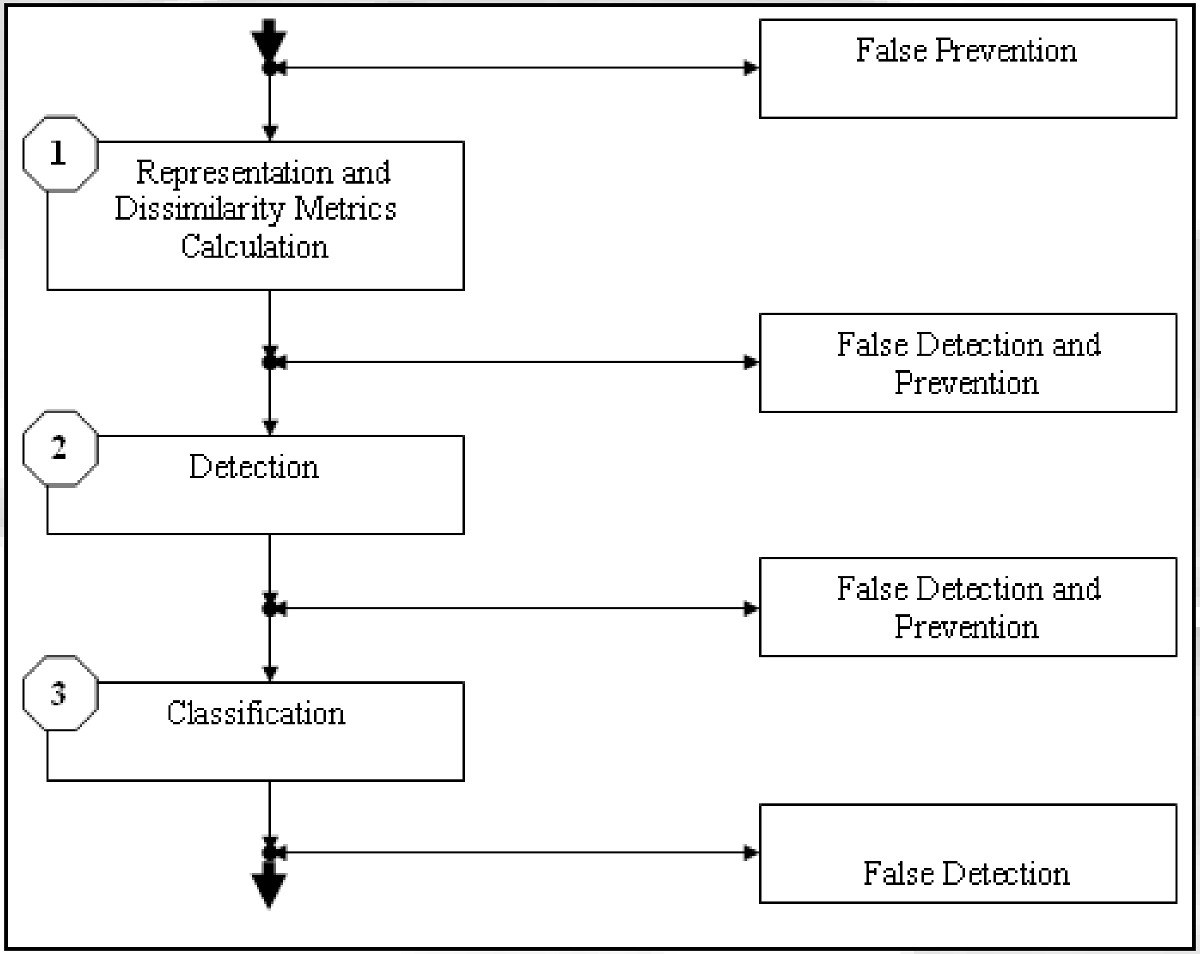


Fig.P.1. Illustrates a general process flow for temporal video segmentation algorithms.

Problems & Solutions

- (P) Quality of Detection
  - (S) Parallel Analyzer
  - (S) Uncertainty Groups
  - (S) Extremely Sensitive Change Detector (ESCD)
  - (S) False Negative Prevention (No Threshold)
  - (S) False Positive Detection
- (P) Complexity vs. Simplicity
  - (S) Example based Technique
  - (S) Uncertainty Groups
- (P) Real Time
  - (S) Adaptive Examples
- (P) Generality vs. Specificity
  - (S) Example based Technique
- (P) Flexibility and Extensibility
  - (S) Multi-level property
  - (S) Example based Technique

Outline

1. Overview
2. Literature Review
3. **Segmentation based on Predefined Examples (S.P.E.)**
4. Segmentation based on Adaptive Examples (S.A.E.)
5. Experimentation Results
6. Conclusions

3. Segmentation based on Predefined Examples (S.P.E.)

3.1. Representation

3.2. Detection

3.3. Classification

3.1. Representation

- **Predefined Examples**
  - Quality
    - Examples of different durations
    - Different types of examples
      - Transition Types
      - Video Types
    - Balance
    - Color Variety
    - Combined Transitions

  - Quantity

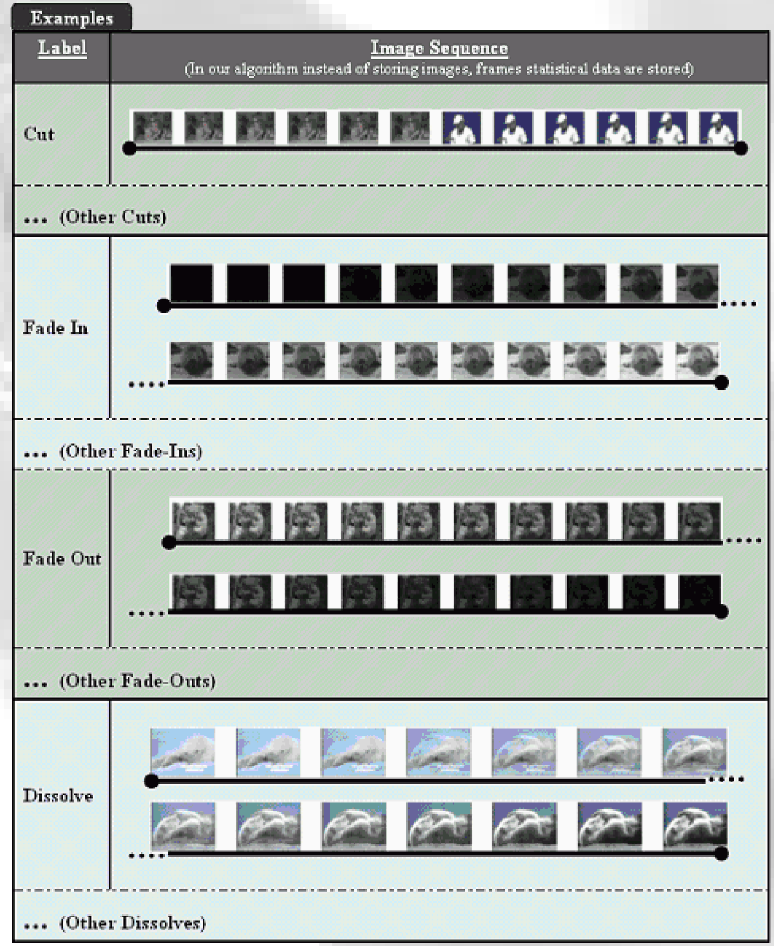


Fig. P.2. Sample cut, fade in, fade out, and dissolve sequences.

3.1. Representation

- **Color Moments (Statistics)**

Statistics \ Color Channels:	Red	Green	Blue
Color Intensities			
Mean:	M_R	M_G	M_B
Standard deviation:	S_R	S_G	S_B
Skew:	K_R	K_G	K_B
Color Center of Gravity			
Horizontal Position			
Mean:	M_{xR}	M_{xG}	M_{xB}
Standard deviation:	S_{xR}	S_{xG}	S_{xB}
Skew:	K_{xR}	K_{xG}	K_{xB}
Vertical Position			
Mean:	M_{yR}	M_{yG}	M_{yB}
Standard deviation:	S_{yR}	S_{yG}	S_{yB}
Skew:	K_{yR}	K_{yG}	K_{yB}

Table P.1. Organizes the twenty seven moments in an easy to understand fashion

Fig. P.3. Illustrates the center of gravities for each of the three color components in real life picture.

3.1. Representation

$$M(t,c) = \frac{1}{N}\sum_{xy} I(x,y,t,c) \qquad \text{(P.1)}$$

$$S(t,c) = \sqrt{\frac{1}{N}\sum_{xy}\left[I(x,y,t,c) - M(t,c)\right]^2} \qquad \text{(P.2)}$$

$$K(t,c) = \sqrt[3]{\frac{1}{N}\sum_{xy}\left[I(x,y,t,c) - M(t,c)\right]^3} \qquad \text{(P.3)}$$

$$M_x(t,c) = \frac{1}{N}\sum_{xy}\frac{I(x,y,t,c)\cdot x}{M(t,c)} \qquad \text{(P.4)}$$

$$S_x(t,c) = \sqrt{\frac{1}{N\cdot M_x(t,c)}\sum_{xy}\left[I(x,y,t,c)\cdot\left(x - M_x(t,c)\right)^2\right]} \qquad \text{(P.5)}$$

$$K_x(t,c) = \sqrt[3]{\frac{1}{N\cdot M_x(t,c)}\sum_{xy}\left[I(x,y,t,c)\cdot(x - M_x(t,c)^3\right]} \qquad \text{(P.6)}$$

3.1. Representation

- **Measure of Difference**

$$Diff = \frac{\alpha \cdot \left[\sum_{f=0}^{F} \left[\frac{\sum_{i=0}^{I} \left| m'_{f,i} - M'_{f,i} \right|^{p}}{I} \right]^{1/p} \right]}{F} + \frac{\beta \cdot \left[\sum_{f=0}^{F} \left[\frac{\sum_{i=0}^{I} \left| d'_{f,i} - D'_{f,i} \right|^{p}}{I} \right]^{1/p} \right]}{F} \quad \text{(P.7)}$$

where

$m_{f,i} = \text{ moment } i \text{ of the current frame from the input stream}$

$d_{f,i} = \text{ derivative of } m_{f,i}$

$M_{f,i} = \text{ moment } i \text{ of current frame from the current example}$

$D_{f,i} = \text{ derivative of } M_{f,i}$

$m'_{f,i} = m_{f,i} \cdot w_i$

$w_i = \text{ weight used for moment } i$

$d'_{f,i} = d_{f,i} \cdot w_i = \text{derivative of } m'_{f,i}$

$M'_{f,i} = M_{f,i} \cdot w_i$

$D'_{f,i} = D_{f,i} \cdot w_i = \text{derivative of } M'_{f,i}$

3.1. Representation

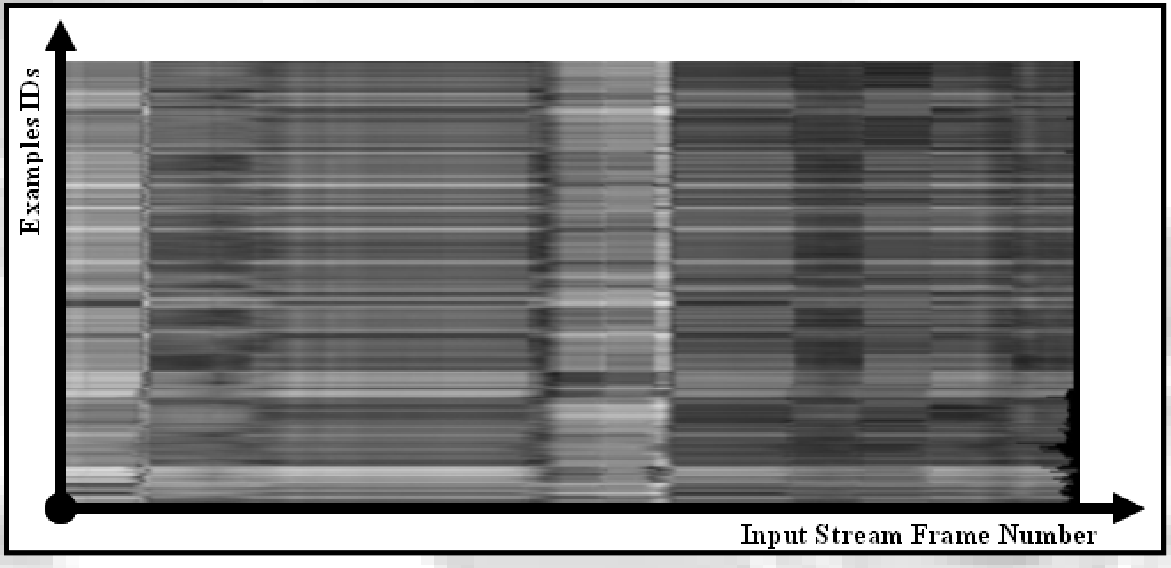


Fig. P.4. Illustrates generated fit values image.

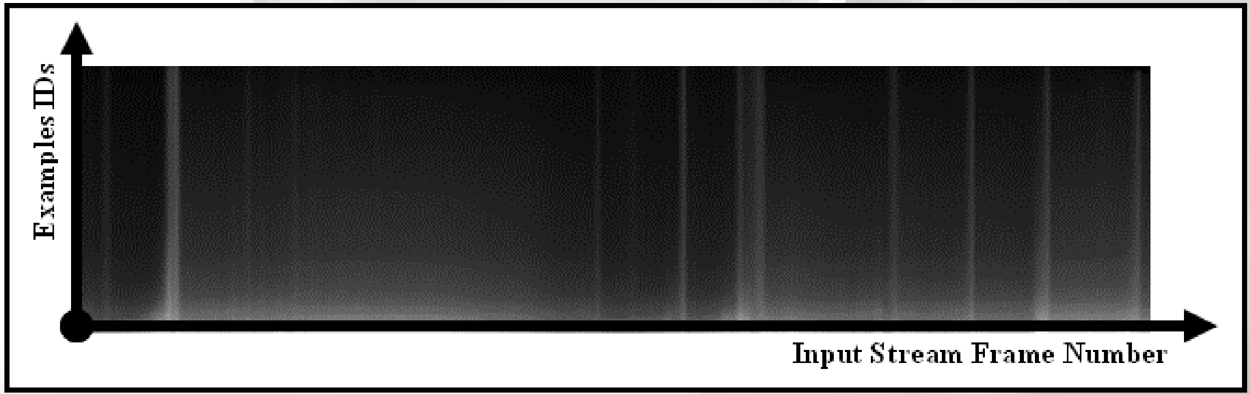


Fig. P.5. Illustrates sorted fit values image

3.2. Detection
3.3. Classification

- Best Examples Extraction & Labeling

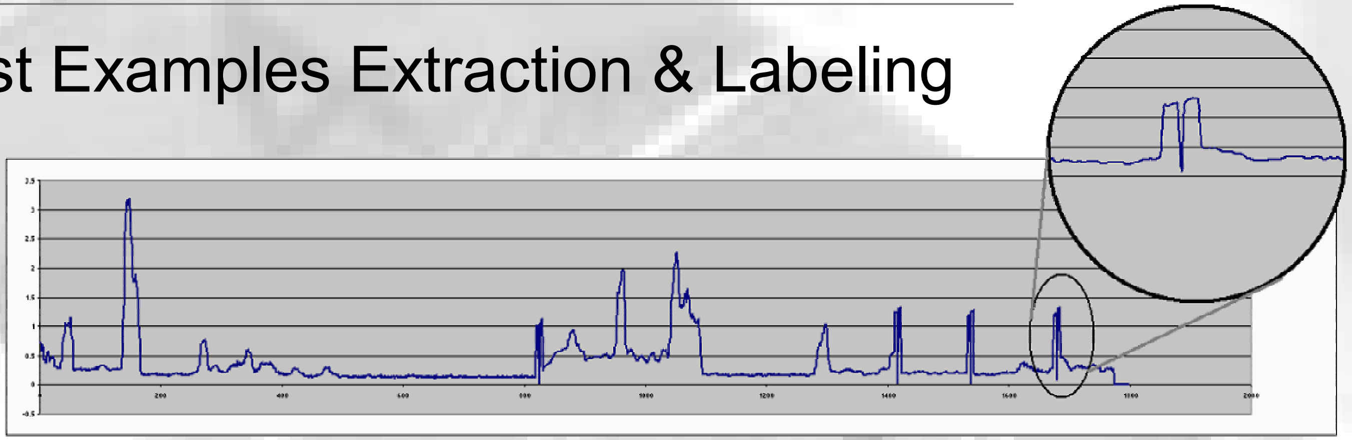


Fig. P.6. Represents the best fit values for each window for one minute of input data.

- Localized Adaptive Threshold

$$threshold = \overline{m}_{f,i} + K \cdot \sigma_{f,i} \qquad \text{(P.8)}$$

Outline

1. Overview
2. Literature Review
3. Segmentation based on Predefined Examples (S.P.E.)
4. **Segmentation based on Adaptive Examples (S.A.E.)**
5. Experimentation Results
6. Conclusions

4. Segmentation based on Adaptive Examples (S.A.E.)

4.1. Representation

4.2. Detection

4.3. Classification

4.4. False Detection and Prevention

4.1. Representation

- **Color Moments (Statistics)**
  - Refer to S.P.E. representation section.

4.1. Representation

- **Adaptive Examples**
  - Cut
  - Dissolve
  - Fade
  - Normal Groups

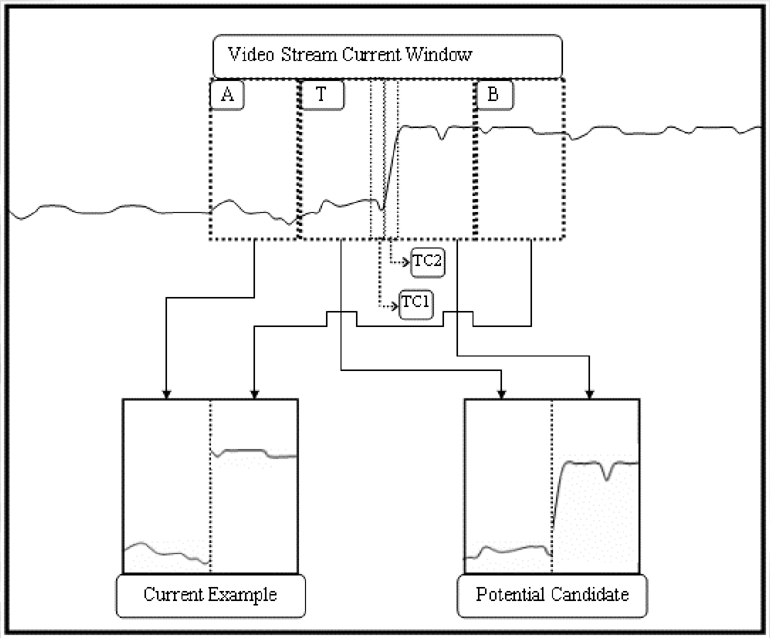


Fig. P.7. Illustrates the process of extracting potential candidate and generating a cut adaptive example while T partition of the window is centered on a cut transition.

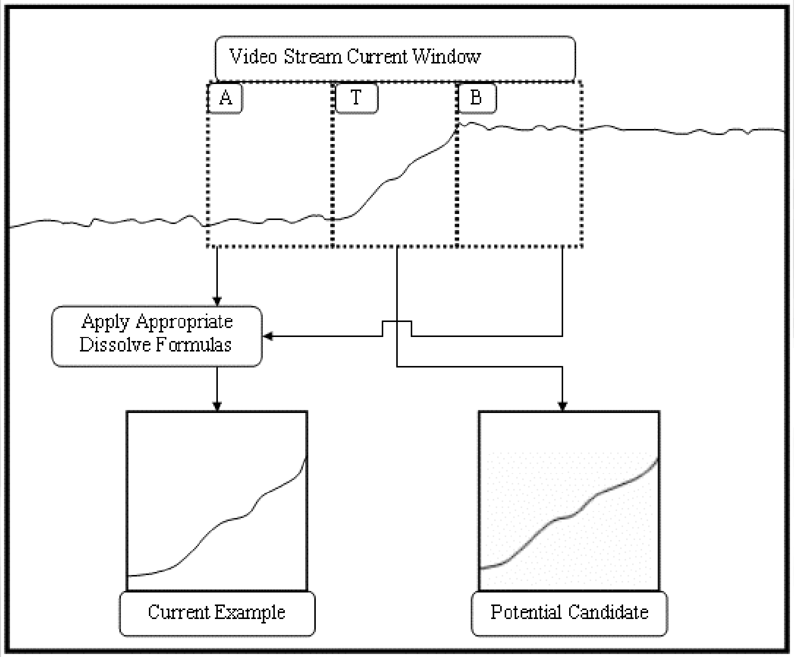


Fig. P.8. Illustrates the process of extracting potential candidate and generating a dissolve adaptive example while T partition of the window is centered on a dissolve transitions.

4.1. Representation

- **Dissolve & Fade (Adaptive Examples)**

$$M_{t,i,R} = \alpha \cdot \overline{A}_{t,i,R} + (1-\alpha) \cdot \overline{B}_{t,i,R} \qquad \text{(P.9)}$$

$$\sigma_{t,i,R} \approx \sqrt{\alpha^2 \cdot \sigma^2_{A_{t,i,R}} + (1-\alpha)^2 \cdot \sigma^2_{B_{t,i,R}}} \qquad \text{(P.10)}$$

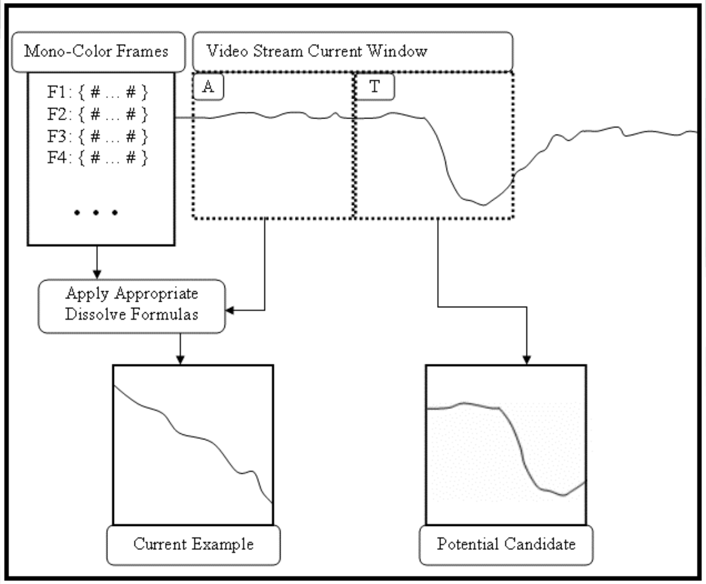


Fig. P.9. Illustrates the process of extracting potential candidate and generating a fade in adaptive example while T partition of the window is centered on a fade in transition.

4.1. Representation

- **Normal Groups (Adaptive Examples)**
  - No Threshold
  - Extremely Sensitive Change Detector

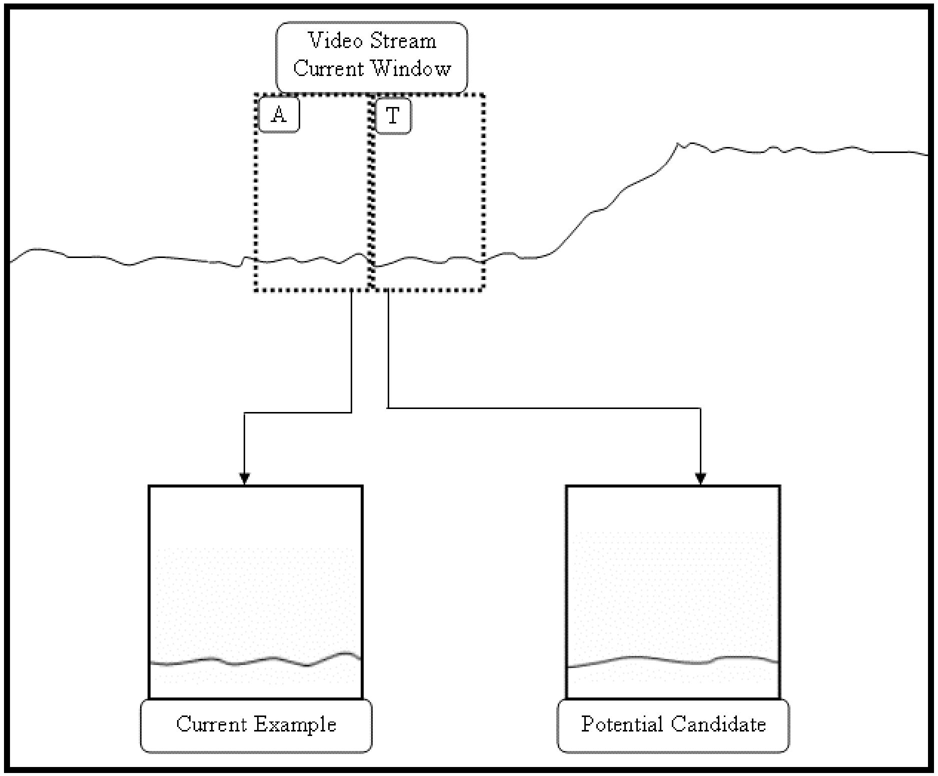
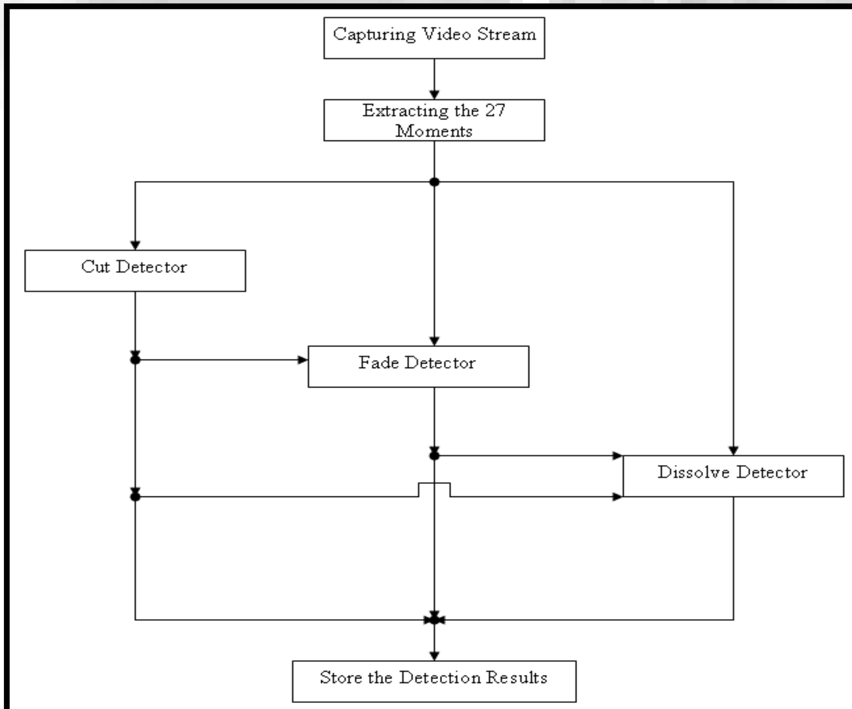


Fig. P.10. Illustrates the process of extracting potential candidate and generating a normal adaptive example for gradual transitions detector while T partition of the window is over a region of no activity (regions containing minor object motions).

4.2. Detection
4.3. Classification
4.4. False Prevention & Detection

- Best Examples Extraction & Labeling
- No Threshold
-



- Parallel Analyzer
- Uncertainty Groups Analyzers
- False Detection Techniques

Fig. P.11. Illustrates the high level process flow for the second algorithm.

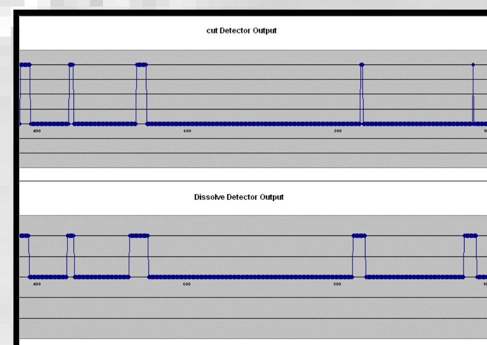


Fig. P.12. Illustrates cut and dissolve detection streams.

Outline

1. Overview
2. Literature Review
3. Segmentation based on Predefined Examples (S.P.E.)
4. Segmentation based on Adaptive Examples (S.A.E.)
5. Experimentation Results
6. Conclusions

5. Experimentation Results

5.1. Evaluation Techniques

- Manual Detection (Truth Data)
- True Positives
- False Positives
- False Negatives
- True Negatives
- Recall
- Precision
- Utility

5.1. Evaluation Techniques

- Recall

$$\mathrm{Re}\,call = R^x = \frac{N^x_{correct}}{N^x_{correct} + N^x_{missed}} \times 100\% \qquad \text{(P.11)}$$

where

$$TP = N^x_{correct} = |\Theta|, \quad \Theta = \{S^x_i, i \in \{1,\dots,k^x_a\} | \exists j \in \{1,\dots,k^x_m\} \ and \ S^x_i \cap S^x_j \neq \phi\}$$

$$FN = N^x_{missed} = |\Theta|, \quad \Theta = \{S^x_i, i \in \{1,\dots,k^x_a\} | \forall j \in \{1,\dots,k^x_m\} \ and \ S^x_i \cap S^x_j = \phi\}$$

5.1. Evaluation Techniques

- Precision

$$\Pr ecision = P^x = \frac{N^x_{correct}}{N^x_{correct} + N^x_{false}} \times 100\% \qquad \text{(P.12)}$$

where

$$TP = N^x_{correct} = |\Theta|, \quad \Theta = \{S^x_i, i \in \{1,\dots,k^x_a\} | \exists j \in \{1,\dots,k^x_m\} \ and \ S^x_i \cap S^x_j \neq \phi\}$$

$$FP = N^x_{false} = |\Theta|, \quad \Theta = \{S^x_j, j \in \{1,\dots,k^x_m\} | \forall i \in \{1,\dots,k^x_a\} \ and \ S^x_i \cap S^x_j = \phi\}$$

5.1. Evaluation Techniques

- Utility
  - General Definition

$$Utility = \alpha \cdot Recall + (1-\alpha) \cdot Precision \quad \text{(P.13)}$$

  - Variation Used

$$Utility = \frac{(Recall + Precision)}{2} \quad \text{(P.14)}$$

5.2. S.P.E. Results

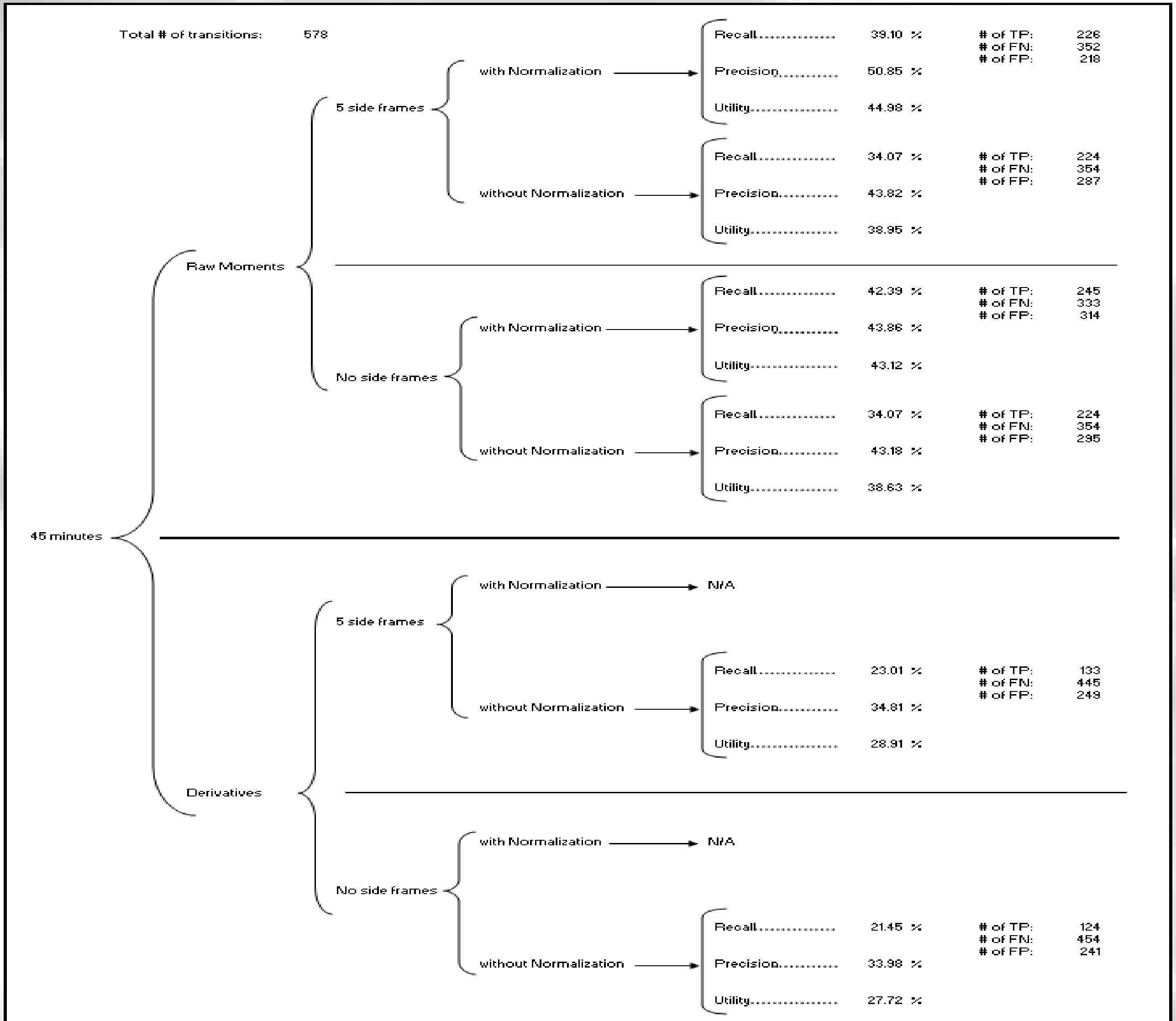


Fig. P.13. Presents the experimentation results for 45 minutes of data.

5.2. S.P.E. Results

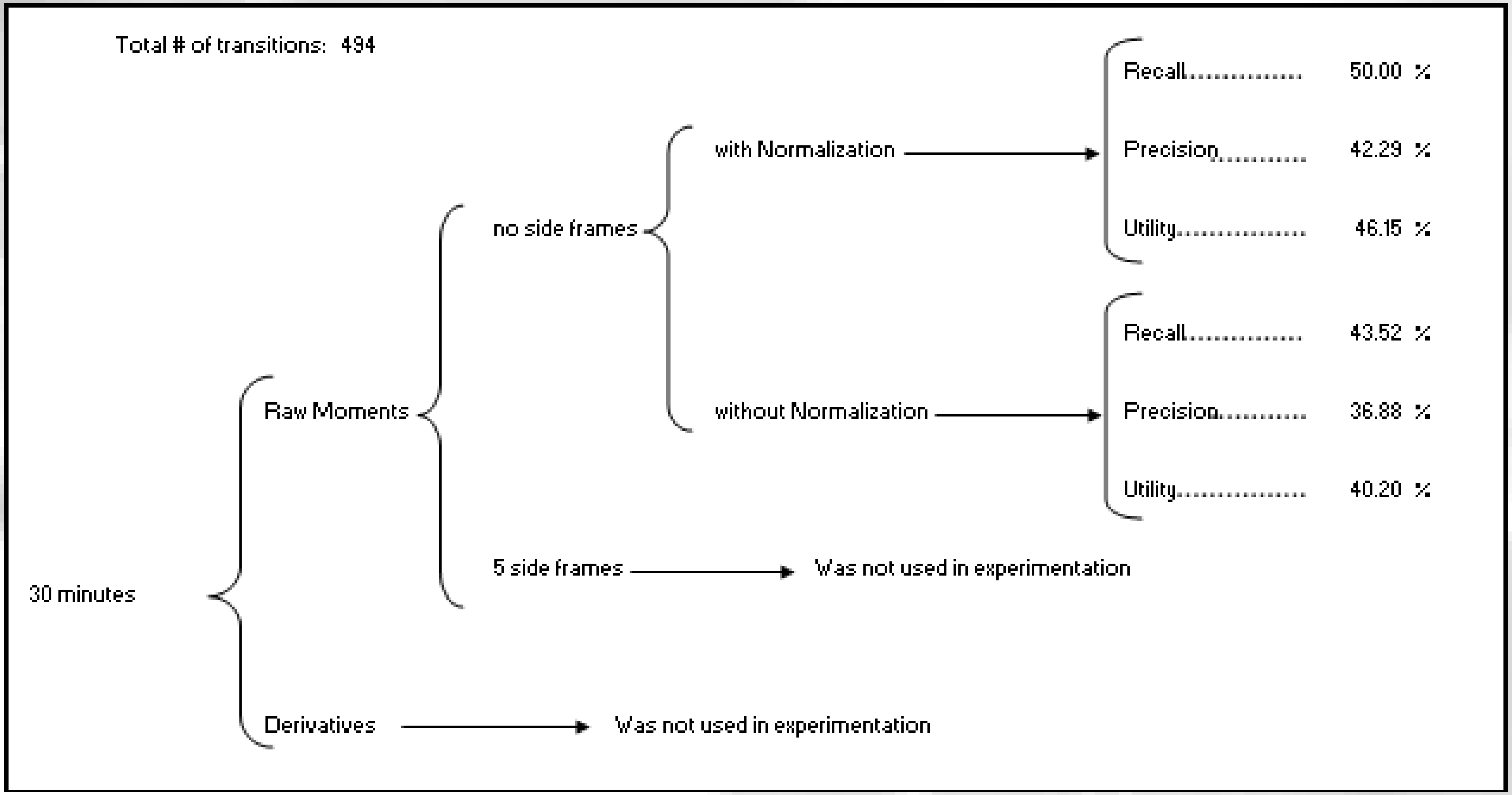


Fig. P.14. Presents the experimentation results for 30 minutes of data.

5.3. S.A.E. Results

	Match (True Positives)	False Alarm (False Positives)	Missed (False Negative)	Recall	Precision	Utility
Cuts	578	20	34	94.44%	96.66%	95.55%
Fades	41	3	0	100.00%	93.18%	96.59%
Dissolves	57	40	3	95.00%	58.76%	76.88%
Total	676	63	37	94.81%	91.47%	93.14%

Table P.2. Presents the final results of the second algorithm.

5.3. S.A.E. Results

Threshold	TP	FP	FN	Recall	Precision	Utility
1.20	599	95	13	97.88%	86.31%	92.93%
1.40	599	95	13	97.88%	86.31%	92.93%
1.60	599	95	13	97.88%	86.31%	92.93%
1.80	599	95	13	97.88%	86.31%	92.93%
2.00	599	95	13	97.88%	87.52%	92.93%
2.20	596	85	16	97.39%	87.52%	92.45%
2.40	594	72	18	97.58%	89.19%	93.12%
2.60	593	63	19	96.90%	90.40%	93.65%
2.80	590	56	22	96.41%	91.33%	93.87%
3.00	590	51	22	96.41%	92.43%	94.22%
3.20	590	48	22	96.41%	92.48%	94.44%
3.40	590	41	22	96.50%	93.50%	94.95%
3.60	585	37	27	95.59%	94.51%	94.82%
3.80	582	31	30	95.42%	94.35%	95.20%
4.00	584	35	28	95.98%	94.94%	94.88%
4.20	580	24	32	94.77%	96.26%	95.40%
4.40	578	22	34	94.44%	96.33%	95.39%
4.60	578	20	34	94.44%	96.66%	95.55%
4.80	577	19	35	94.28%	96.81%	95.55%
5.00	576	19	36	94.12%	96.81%	95.46%
5.20	573	19	39	93.63%	96.79%	95.21%
5.40	573	17	39	93.63%	97.12%	95.37%
5.60	572	16	40	93.46%	97.28%	95.37%
5.80	569	16	43	92.97%	97.26%	95.26%
6.00	567	15	45	92.65%	97.42%	95.34%
6.20	566	15	46	92.48%	97.42%	94.95%
6.40	563	14	49	91.99%	97.57%	94.78%
6.60	562	12	50	91.83%	97.91%	94.87%
6.80	560	12	52	91.50%	97.90%	94.70%
7.00	554	12	58	90.52%	97.88%	94.20%

Table P.3. Presents number of true positives, false negatives, false positives, as well as recall, precision and utility for different thresholds used in false positive detector of cut detector.

5.3. S.A.E. Results

ROC Curve

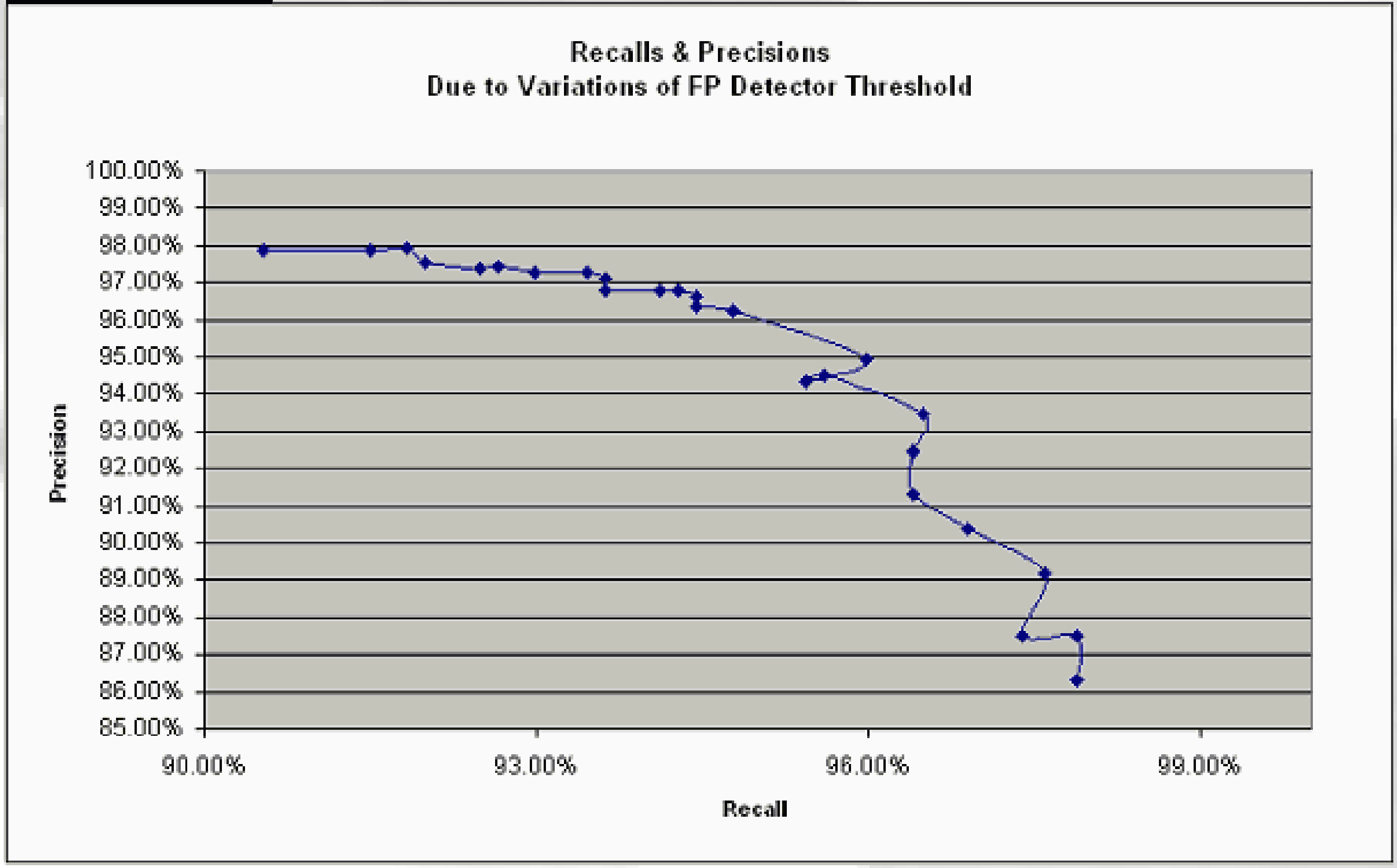


Fig. P.15. Presents the recall and precision values for different thresholds used in false positive detector of cut detector as well as the ROC curve for the second algorithm.

5.3. S.A.E. Results

Utility Curve

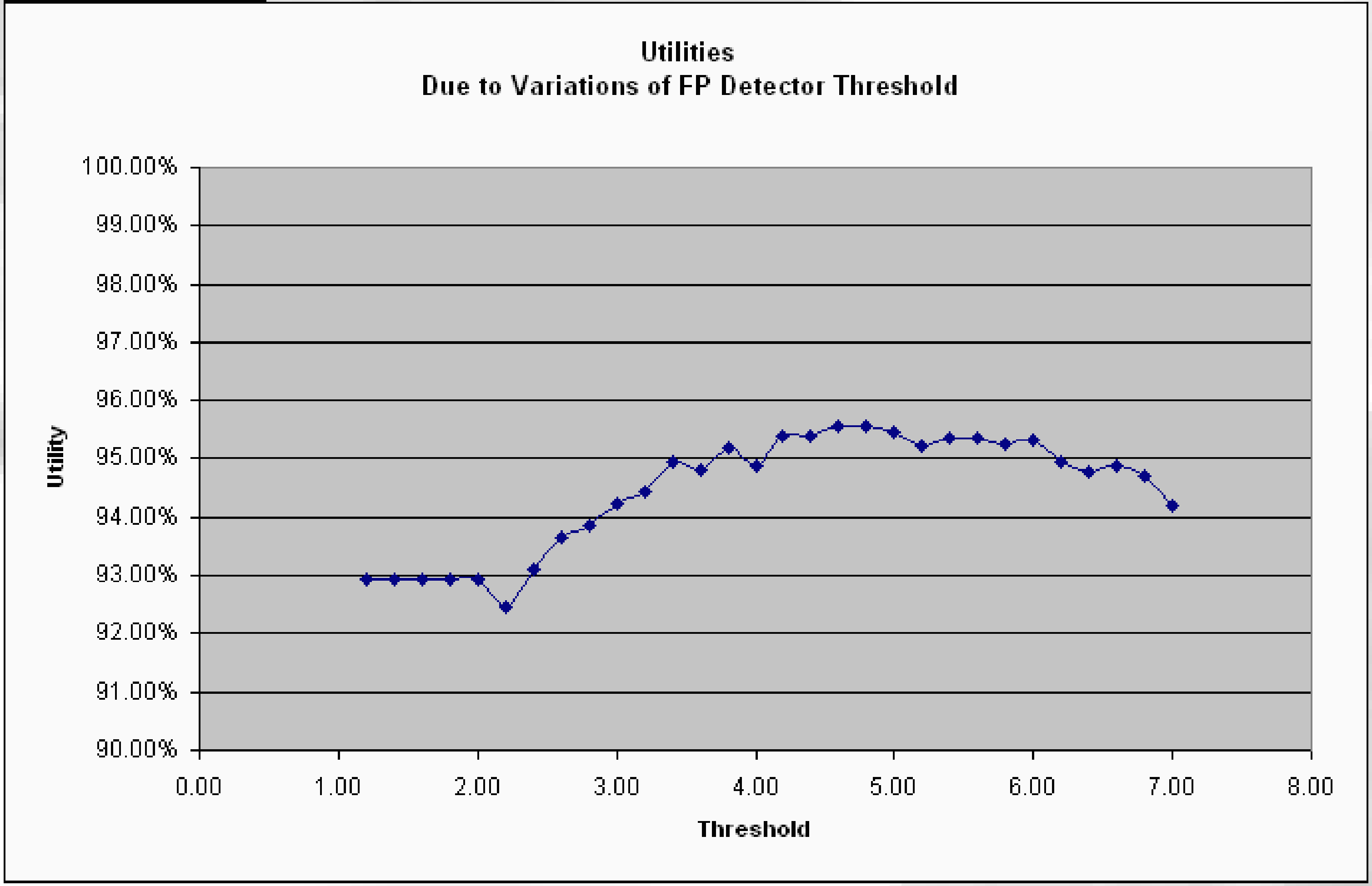


Fig. P.16. Presents the utility values for different thresholds used in false positive detector of cut detector as well as the utility curve for the second algorithm.

5.4. Discussion

- S.P.E.
  - \+ Simplicity
  - \+ Generality
  - \+ Flexibility
  - \+ Extensibility
  - \- Real Time
  - \- High Quality Detection

- S.A.E.
  - \+ Simplicity
  - ~ Generality
  - \+ Flexibility
  - ~ Extensibility
  - \+ Real Time
  - \+ High Quality Detection

Table P.4. Presents the time performance of the second algorithm for one minute of data.

	Execution Time
Statistical Data Preparation	1.0 seconds
Cuts	2.1 seconds
Fades	10.8 seconds
Dissolves	0.3 seconds
Total	14.2 seconds

Outline

1. Overview
2. Literature Review
3. Segmentation based on Predefined Examples (S.P.E.)
4. Segmentation based on Adaptive Examples (S.A.E.)
5. Experimentation Results
6. Conclusions

6. Conclusions

6.1. Summary

- Two methods were implemented and tested
  - The first one was based on lots of *predefined* examples
  - The second one was based on *adaptive* examples
- The latter method outperformed the first
- Our solutions directed all the problems of previous works:
  - High Quality Detection
  - Simplicity
  - Real time
  - Generality
  - Flexibility
  - Extensibility

6.2. Future Works

- Future Enhancements
  - Direct Comparison based on Predefined Examples
  - Direct Comparison based on Adaptive Examples
- Next Generation Algorithm
  - Simultaneous detections & use of specialized clustering methods to increase generality

6.3. Acknowledgements

Dr. John Gauch (Thesis Committee Chair)

Dr. Arvin Agah (Thesis Committee Member)

Dr. James Miller (Thesis Committee Member)

6.4. References

1) Yeganeh, Robert, "A Multi-Level Approach for Video Shot Boundary Detection based on Adaptive Examples," M.S. Thesis, KU, Lawrence, KS, 2006.

2) *Refer to [1] for list of references used during our work.*

6.5. Q & A Session