EG2013 Tutorial on VIDEO VISUALIZATION

7. Applications of Video Visualization

Phil Legg
Swansea University

What applications does it suit?

Video Visualization serves as a summary tool for quick review and understanding of video data.

- Can the Video Visualization be interpreted faster than watching the video?
- Is the application time-dependent?
- Is there too much video data to reasonably expect a human to have to watch?
- Can Video Visualization indicate data trends that may not actually be recognised by a human viewing?
- Are there other scenarios where Video Visualization may be more useful or may complement the original video?

Types of application

In the literature, there are five main application areas where video visualization has been used.

- Surveillance video
- Television programmes.
- Entertainment
- Sport
- Facial expressions

Other potential applications for the future could also include medical visualization, product manufacturing, and military video analysis.

Surveillance

- Extremely vast amount of video data collected
  - could be as much as 24 hours every day.
- However, much of this recorded video may be unimportant or irrelevant – for example, a static scene.
- The operator needs to determine moments of importance from this large data – e.g., crime (fighting, robbery, trespassing, etc.)

Video Visualization facilitates rapid review of important segments in video content.

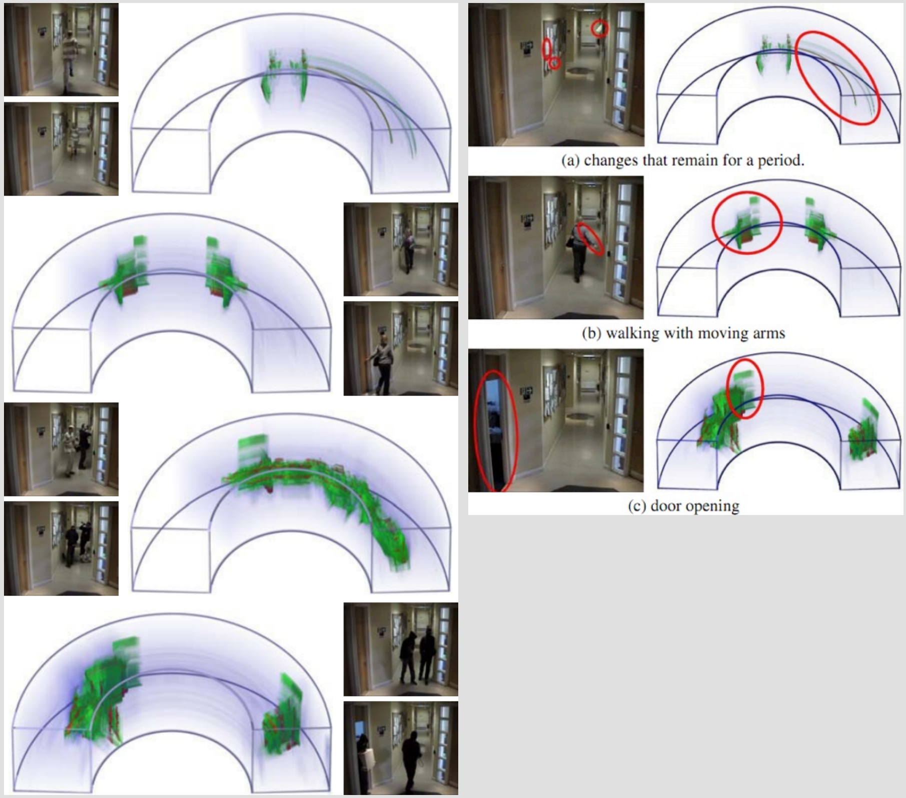
TV Programmes

- Viewer may want to skip to a particular moment in recorded TV footage.
- May not be exactly sure of the time in the video that this particular moment occurs.

Video Visualization facilitates rapid seeking of video content to find relevant segment.

Daniel and Chen 2003, IEEE Vis

Surveillance

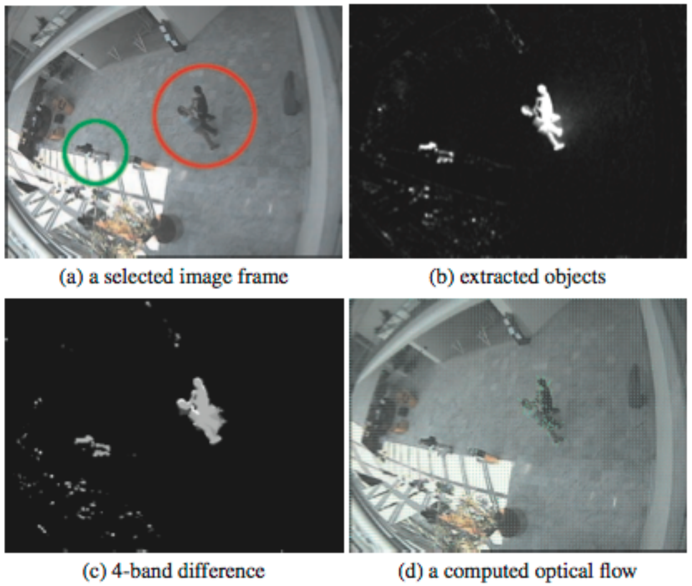


Television

Aim: Video Visualization is a function to create effective visualization images from video. Users can then recognize different spatiotemporal entities 'at once'.

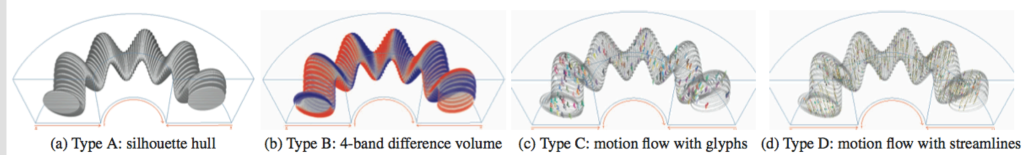
Chen et. al., 2006, IEEE TVCG (Vis)

Surveillance
(Video clips from CAVIAR Project)



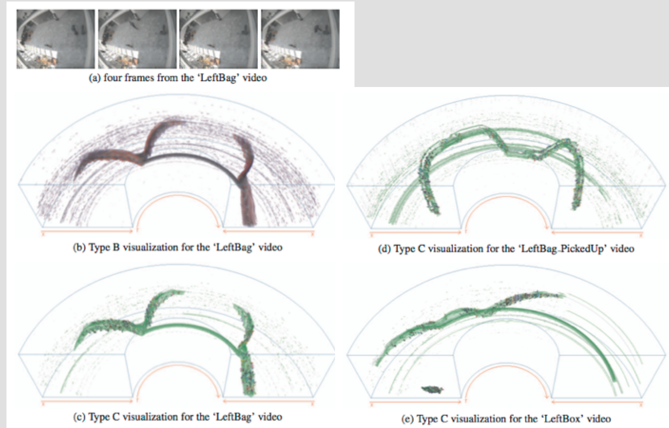
(a) a selected image frame (b) extracted objects (c) 4-band difference (d) a computed optical flow

- Extracting objects from scene
- 4-band difference between video frames (Background, Disappearing, Overlapping and New).
- Optical flow field (motion).

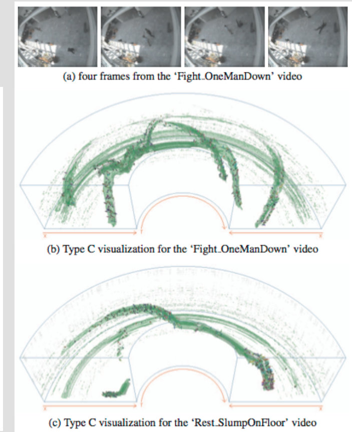


(a) Type A: silhouette hull (b) Type B: 4-band difference volume (c) Type C: motion flow with glyphs (d) Type D: motion flow with streamlines

- 4 visualization approaches trialed in a user study – Type B preferred overall, Type C preferred for spin motion.



(a) four frames from the 'LeftBag' video
(b) Type B visualization for the 'LeftBag' video
(d) Type C visualization for the 'LeftBag_PickedUp' video
(c) Type C visualization for the 'LeftBag' video
(e) Type C visualization for the 'LeftBox' video

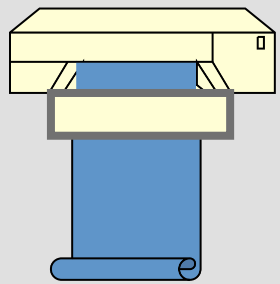


(a) four frames from the 'Fight_OneManDown' video
(b) Type C visualization for the 'Fight_OneManDown' video
(c) Type C visualization for the 'Rest_SlumpOnFloor' video

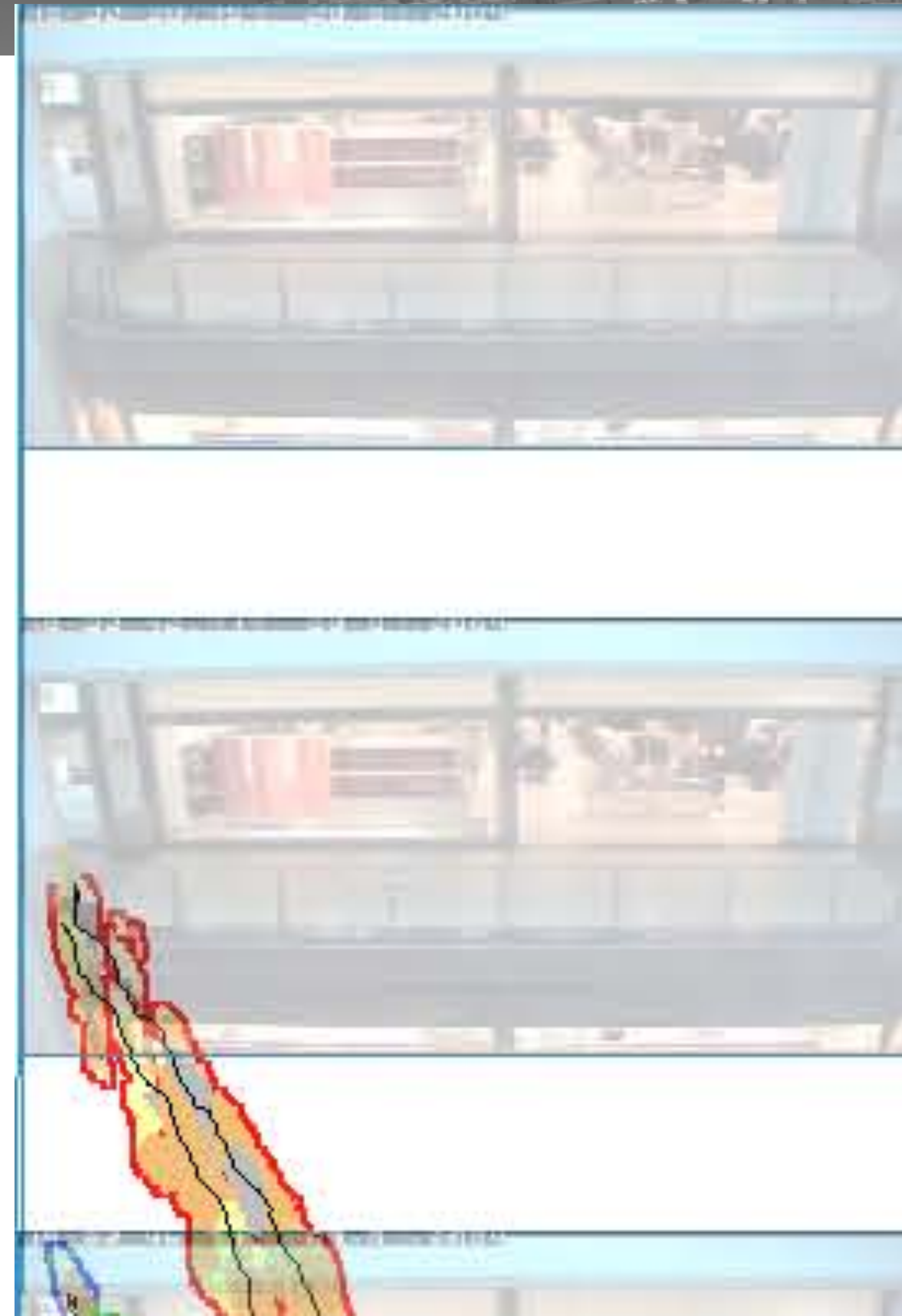
- Left: Four example videos of leaving objects in the scene.
  - Note the static paths of stationary objects.
  - (c) owner leaves the scene then reappears.
  - (d) object was left for brief period, owner not far away.
  - (e) object left for long period of time, owner walks away.
- Right: Two more example videos of the scene.
  - Static path and lack of arrow glyphs suggests little motion whilst actor on the floor.

Botchen et. al., 2008, IEEE TVCG

VPG (VideoPerpetuoGram)



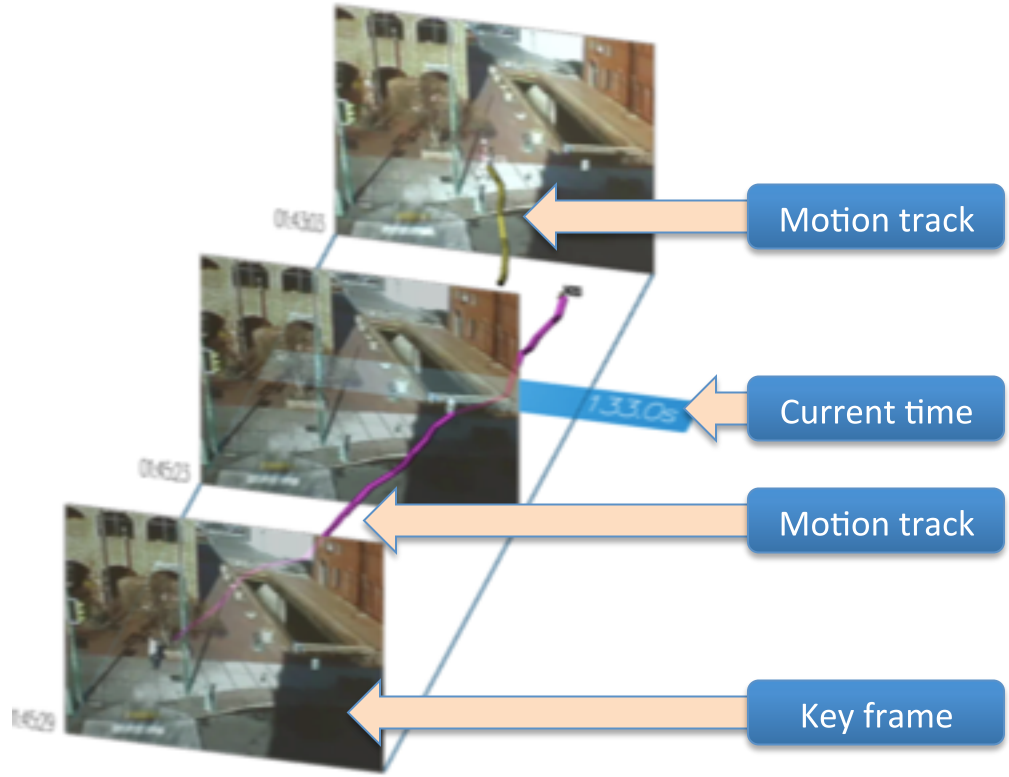
Record motion similar to an Electrocardiogram (ECG) and Seismographs.



Bosch et. al., 2009, IEEE VAST

Surveillance

- Identify the encounter of people by their movement trajectories.
- Optical flow and background subtraction.
- Camera calibration used to calculate object properties (mean speed, average direction, perspective-corrected viewpoints also computed).



Höferlin et. al., 2012, IEEE TVCG (Vis)

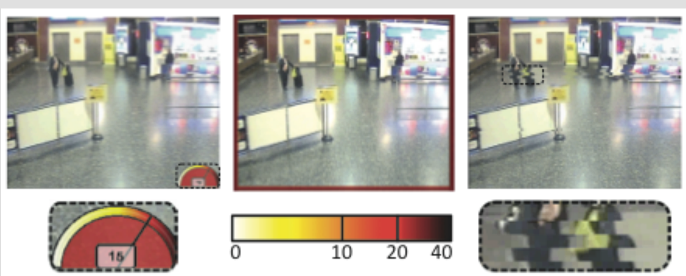
Fast Forward Video Visualization

Two goals / trade-offs:

- (1) Object identification
- (2) motion perception.

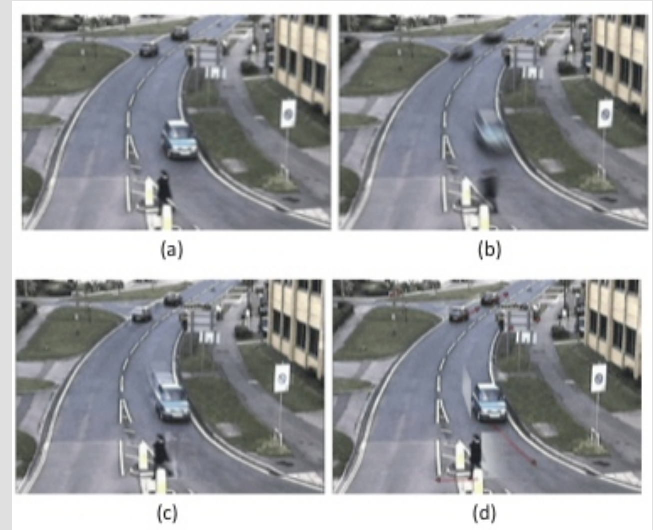
Adaptive Fast Forward Playback Speed Visualization

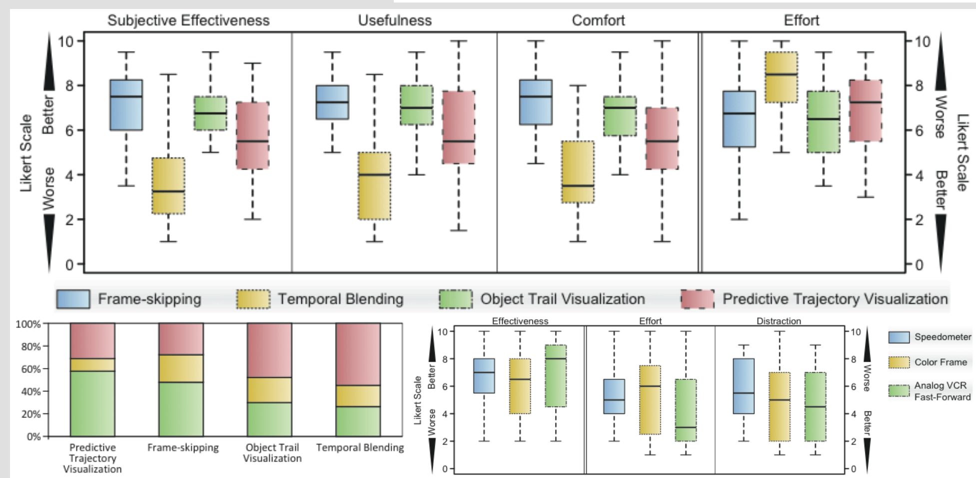
Previous user study suggested to “add visual feedback to increase awareness of playback speed.”



Speedometer (bottom right of video).
Color frame (border of video).
Analog VCR (speed mapped to horizontal distortion).

Four techniques: (a) Frame skipping, (b) Temporal blending, *(c) Object trail ->frame blending and object enhancement, (d) Predictive trajectory -> frame-skipping with motion arrow*



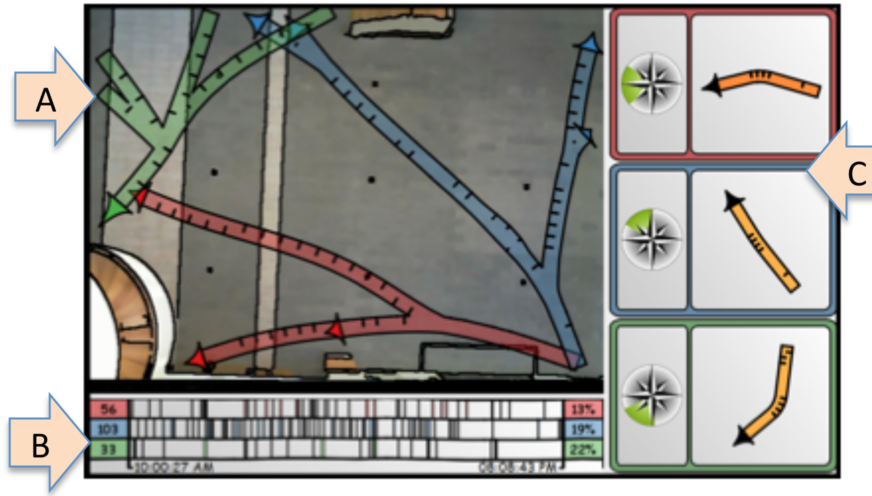


User study favor (a), worse is (b). Motion best observed using (d). Speedometer only just scored highest but feedback suggests that Analog VCR is preferred by participants.

Höferlin et. al., 2013, TMM

Illustrative summary of surveillance video

- (A) Spatial context view indicates major paths using trajectory bundling.
- (B) Temporal context view shows coverage of clusters, number of trajectories and diversity.
- (C) Facet showcase view depicts azimuth coverage, azimuth mean, and cluster velocity.



Entertainment

- Movies consist of many different audio/video components (e.g., background music, sound effects, actors, scenes).
- Those with disabilities may not experience movies in quite the same way to others (e.g., hearing difficulties).

Video Visualization incorporates visual description of multiple audio and video components to convey greater understanding.

Jänicke et. al., 2010, CGF (Eurographics)

SoundRiver: Semantically-Rich Sound Illustration

Visual depiction of movies by mapping components from soundtrack to visual metaphors. Can incorporate additional information such as mood, volume, motion (footsteps), subtitles, number of speakers and gender.

- Motion/footsteps (arrows)
- Volume (height) Mood (colour)
- Audio events (glyphs)
- #speakers, gender, volume

Sport

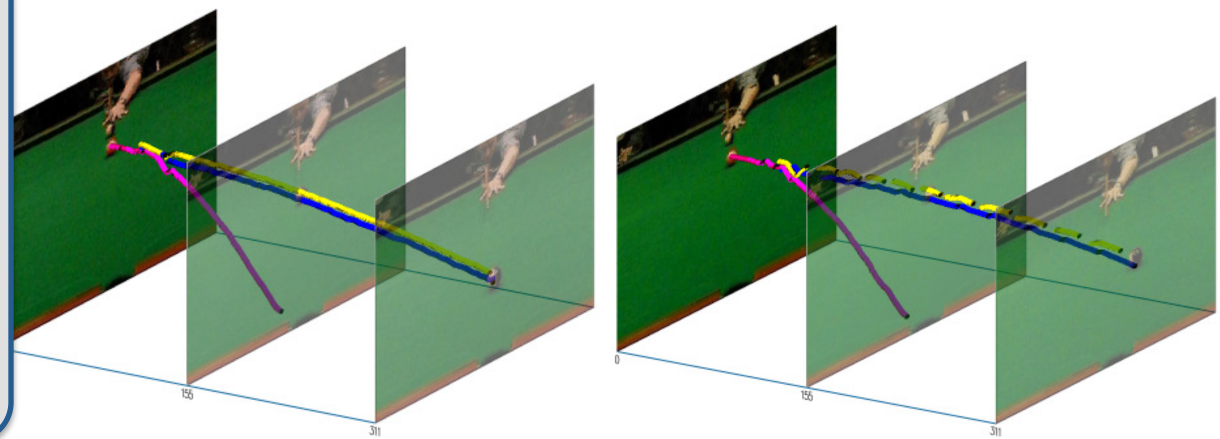
- Players, coaches and analysts all want to recap on particular key elements from a match.
- Video can highlight player and/or team performances.
- Quick decision-making is required during in-match video review.

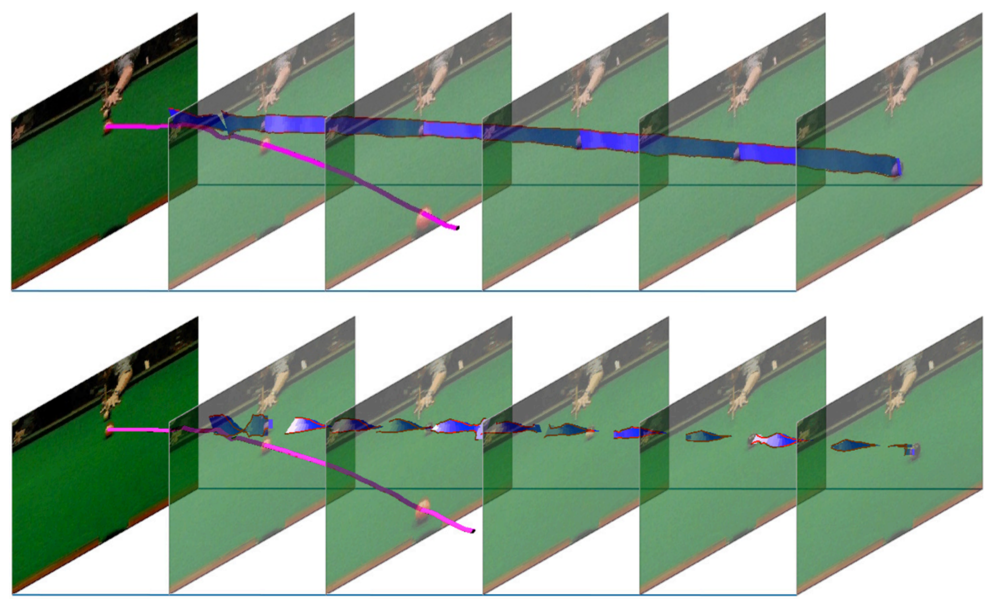
Video Visualization facilitates rapid decision making by highlighting key segments from video.

Höferlin et al., 2012, IEEE TVCG (Vis)

Shot Analysis for training

VPG shows the ball motion between video frames.
Can also depict other shot attributes including ball spin.



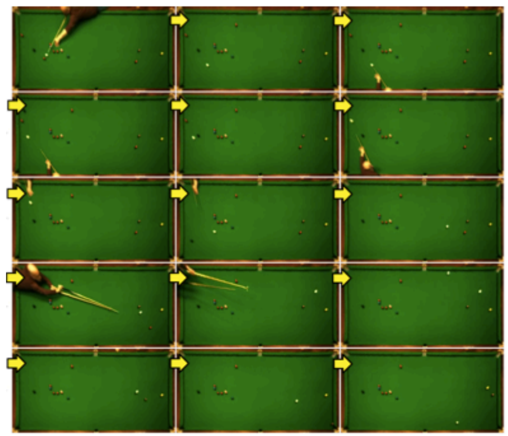


Ribbon-effect to show the spin of the cueball.
Top: Ball played with no side spin creates a flat effect.
Bottom: Ball played with side spin creates a twisted effect.

Parry et. al., 2011, IEEE TVCG (Vis)

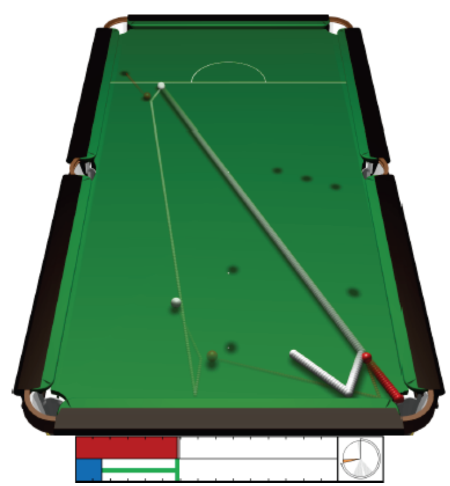
Video storyboard

Illustrates an entire match based on the key events that occurred in that time.
Key shots are emphasized, with preceding and following shots also shown.
Key periods are emphasized by size of illustration.



Dashboard also depicts time in match, player scores (before and after the illustrated period), and remaining points on the table.

Creates a visual summary that coaches can refer to for identifying good and bad periods of play. Could also mean that a coach can train more players at same time.



Legg et. al., 2012, CGF (Eurovis)

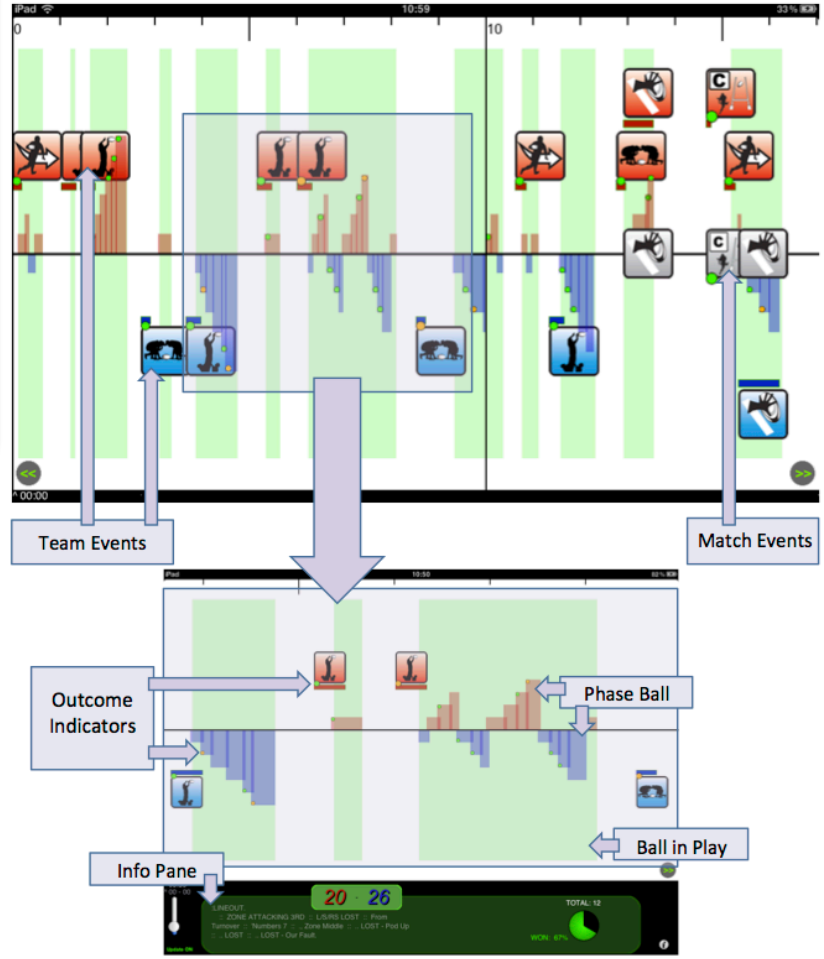
MatchPad: Interactive Glyph-Based Visualization for Real-Time Sports Performance Analysis

- Illustrates match events in real-time based on notational analysis data.
- Provides a graphical overview for rapid recap and decision making during the game.
- Attributes include event, duration, outcome, team.
- Status bar gives detailed view and statistical output.
- iPad-based with direct video playback of events.

The main thing for us is visualizing the data and visualizing it in a very easy to interpret manner

It is a great tool for oversight when matches are very intense, to see what are the key events and how they interact with each other.

With the iPad it is about portability. We have it with us all the time so when the coach wants to know something we can show it immediately on the MatchPad.



Facial Expressions

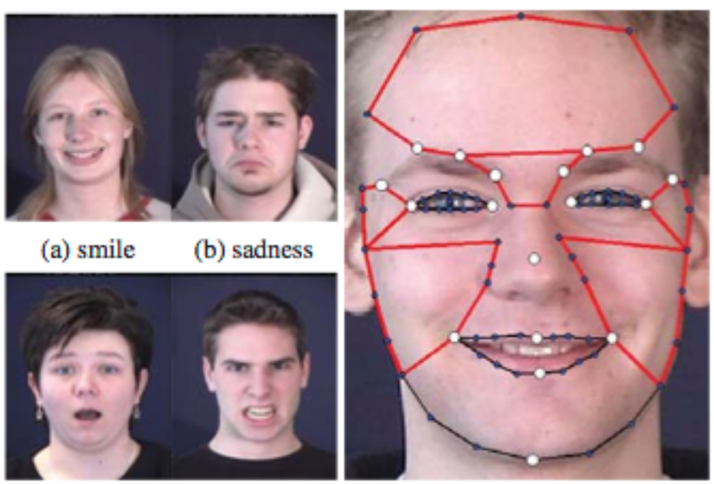
- Video is used for capture of facial expressions for entertainment usage (e.g., animation).

Video Visualization can be used to categorize different characteristics from video data.

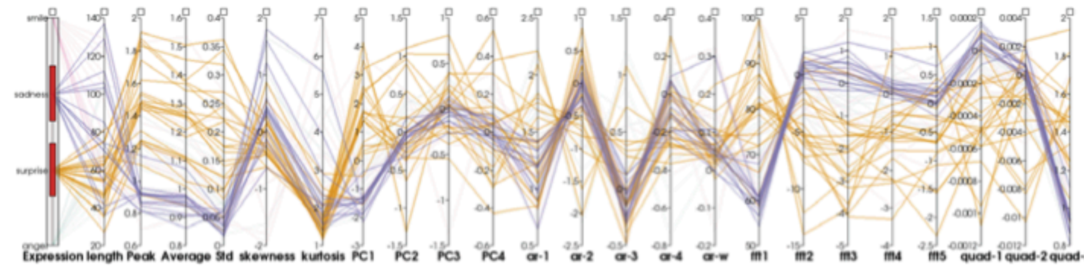
Tam et al., 2011, CGF (EuroVis)

Visualizing facial expressions

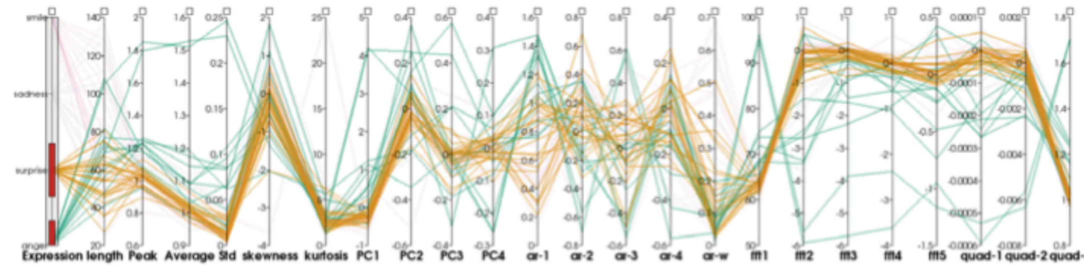
Feature-based time-series data to multi-dimensional parameter space. This enables techniques such as parallel co-ordinates to be used to analyze the data.
Also incorporates analytic view and decision trees.



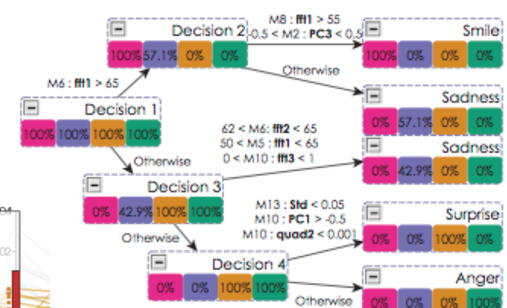
(a) smile (b) sadness (c) surprise (d) anger (e) feature points and regions

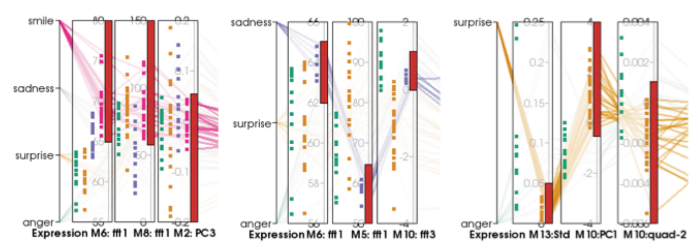


(a) Sadness and Surprise on M5 (mouth height)



(b) Anger and Surprise on M13 (inner brow texture intensity)





(a) Separate smile from all expressions (b) Identifying sadness at Decision 3 (c) Identifying surprise at Decision 4

Conclusions

- We have presented a variety of applications that benefit from video visualization.
- Video Visualization aids rapid understanding of large video content that is time-consuming to watch.
- Can introduce additional knowledge that may not be recognised through traditional viewing and memorization of video content.
- There are most likely many more application areas that these concepts could easily be applied to.