8.8 Content Repurposing

8.8.1 Fundamentals

8.8.2 Text to Speech and Speech to Text

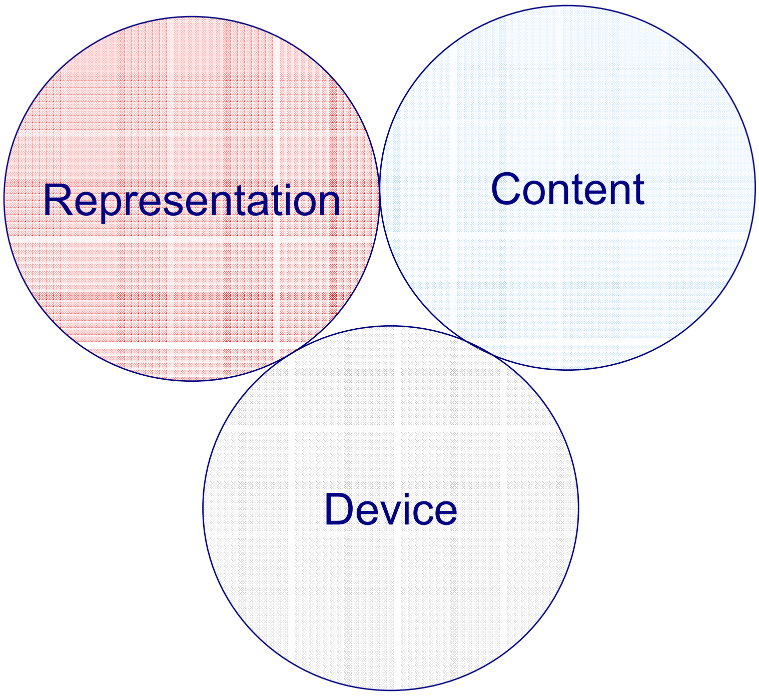
8.8.3 Photo to Video

8.8.4 Video to Panoramic Photo

8.8.5 Repurposing of Panoramic Video: DIVER

Mobile End-User Devices

Re-formatting Content for Specific Devices



Representation

Content

Device

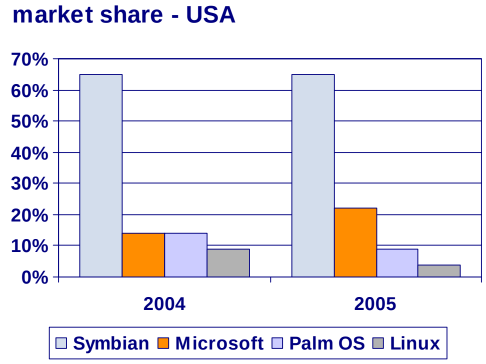
Representations (Formats)

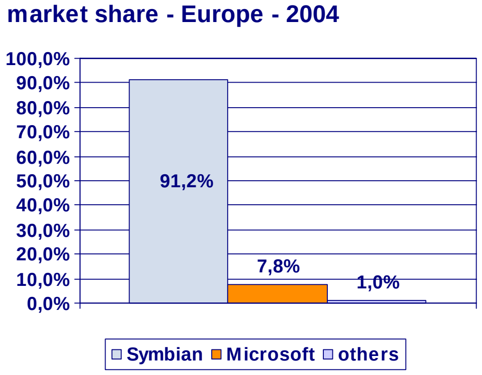
MIME types*)	**number**
audio	62
graphics	36
video	35
text	38
applications	399

*) according to Internet Assigned Numbers Authority (IANA)

Heterogeneous Operating Systems

Example: Smartphones*)





*) Symbian & Smartphone Markt, PDAStreet.com, 03. Juni 2004

8.8.2 Text to Speech and Speech to Text

Text to Speech

Speech generators accept ASCII text as input and produce voice output. They can be used for end-user devices that cannot display text (e.g., classic phones).

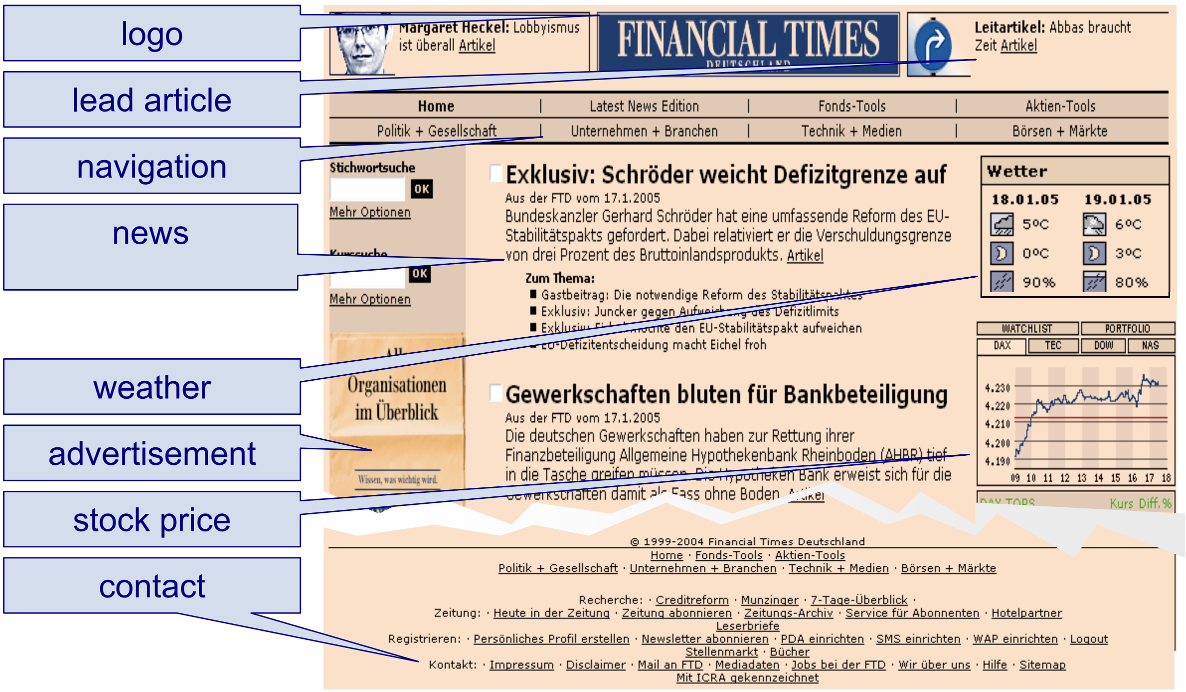
Speech to Text

Dictation systems accept speech as input and produce ASCII output.

Modern dictation software

- no longer requires a training phase by a particular speaker
- can accept continuous speech
- has reasonable word error rates for common text
- can be enhanced with application-specific vocabularies

Transcoding Example



logo

lead article

navigation

news

weather

advertisement

stock price

contact

Margaret Heckel: Lobbyismus ist überall Artikel

FINANCIAL TIMES DEUTSCHLAND

Leitartikel: Abbas braucht Zeit Artikel

Home | Latest News Edition | Fonds-Tools | Aktien-Tools

Politik + Gesellschaft | Unternehmen + Branchen | Technik + Medien | Börsen + Märkte

Stichwortsuche

OK

Mehr Optionen

Kurssuche

OK

Mehr Optionen

Organisationen im Überblick

Wissen, was wichtig wird.

Exklusiv: Schröder weicht Defizitgrenze auf

Aus der FTD vom 17.1.2005

Bundeskanzler Gerhard Schröder hat eine umfassende Reform des EU-Stabilitätspakts gefordert. Dabei relativiert er die Verschuldungsgrenze von drei Prozent des Bruttoinlandsprodukts. Artikel

Zum Thema:

- Gastbeitrag: Die notwendige Reform des Stabilitätspaktes
- Exklusiv: Juncker gegen Aufweichung des Defizitlimits
- Exklusiv: Eichel möchte den EU-Stabilitätspakt aufweichen
- EU-Defizitentscheidung macht Eichel froh

Gewerkschaften bluten für Bankbeteiligung

Aus der FTD vom 17.1.2005

Die deutschen Gewerkschaften haben zur Rettung ihrer Finanzbeteiligung Allgemeine Hypothekenbank Rheinboden (AHBR) tief in die Tasche greifen müssen. Die Hypotheken Bank erweist sich für die Gewerkschaften damit als Fass ohne Boden. Artikel

Wetter	
18.01.05	19.01.05
5°C	6°C
0°C	3°C
90%	80%

WATCHLIST | PORTFOLIO

DAX | TEC | DOW | NAS

DAX TOPS Kurs Diff.%

© 1999-2004 Financial Times Deutschland

Home · Fonds-Tools · Aktien-Tools

Politik + Gesellschaft · Unternehmen + Branchen · Technik + Medien · Börsen + Märkte

Recherche: · Creditreform · Munzinger · 7-Tage-Überblick ·
Zeitung: · Heute in der Zeitung · Zeitung abonnieren · Zeitungs-Archiv · Service für Abonnenten · Hotelpartner
Leserbriefe
Registrieren: · Persönliches Profil erstellen · Newsletter abonnieren · PDA einrichten · SMS einrichten · WAP einrichten · Logout
Stellenmarkt · Bücher
Kontakt: · Impressum · Disclaimer · Mail an FTD · Mediadaten · Jobs bei der FTD · Wir über uns · Hilfe · Sitemap
Mit ICRA gekennzeichnet

Transcoding (1)

Financial Times Deutschland
[Leitartikel]

Aktuelles:
Exklusiv: Schröder weicht Defizitgrenze auf [mehr]
Gewerkschaften bluten für Bankbeteiligungen [mehr]
[weitere]

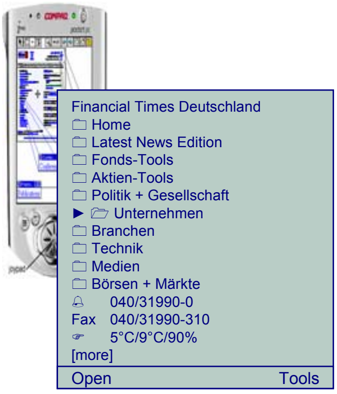
DAX	4.232,36	+0,48
Dow Jones	10.558,00	+0,50
NASDAQ100	1.561,11	+1,03
EUR/USD Depot	1,3104	+0,02

[weitere Kurse]

Financial Times Deutschland

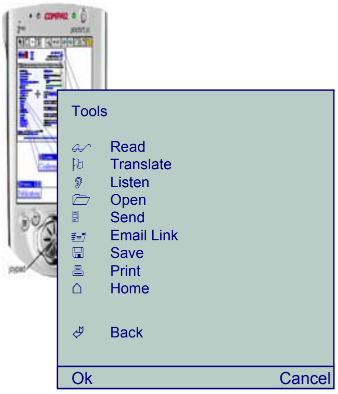
[Image – Heckel_margaret,0.gif]
Magaret Heckel: Lobbyismus ist überall
[Artikel]
[Image - ft_logo_homepage.gif]
[Image - leitartikel.gif]
[Leitartikel: Abbas braucht Zeit Artikel]
[Home] [Latest News Edition]
[Fonds-Tools] [Aktien-Tools]
[Politik + Gesellschaft] [Unternehmen]
[Branchen] [Technik] [Medien]
[Börsen + Märkte]
[Exklusiv: Schröder weicht
Defizitgrenze …]

Transcoding (2)

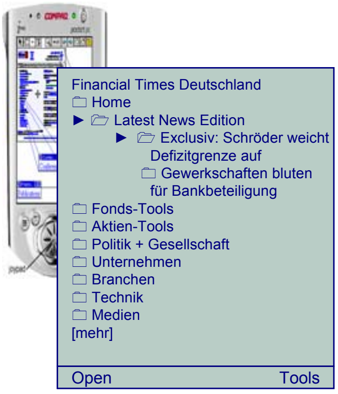


Financial Times Deutschland
- 🗀 Home
- 🗀 Latest News Edition
- 🗀 Fonds-Tools
- 🗀 Aktien-Tools
- 🗀 Politik + Gesellschaft
- ► 🗁 Unternehmen
- 🗀 Branchen
- 🗀 Technik
- 🗀 Medien
- 🗀 Börsen + Märkte

☎ 040/31990-0
Fax 040/31990-310
☞ 5°C/9°C/90%
[more]

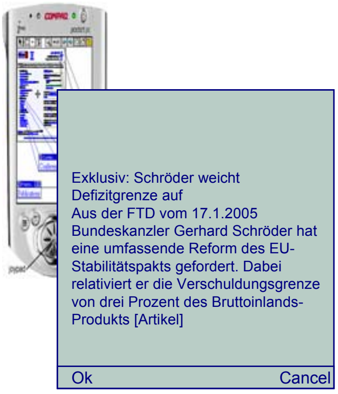


Transcoding (3)



Financial Times Deutschland
- 🗀 Home
- ► 🗁 Latest News Edition
  - ► 🗁 Exclusiv: Schröder weicht Defizitgrenze auf
  - 🗀 Gewerkschaften bluten für Bankbeteiligung
- 🗀 Fonds-Tools
- 🗀 Aktien-Tools
- 🗀 Politik + Gesellschaft
- 🗀 Unternehmen
- 🗀 Branchen
- 🗀 Technik
- 🗀 Medien

[mehr]



Exklusiv: Schröder weicht Defizitgrenze auf
Aus der FTD vom 17.1.2005
Bundeskanzler Gerhard Schröder hat eine umfassende Reform des EU-Stabilitätspakts gefordert. Dabei relativiert er die Verschuldungsgrenze von drei Prozent des Bruttoinlands-Produkts [Artikel]

8.8.3 Photo to Video

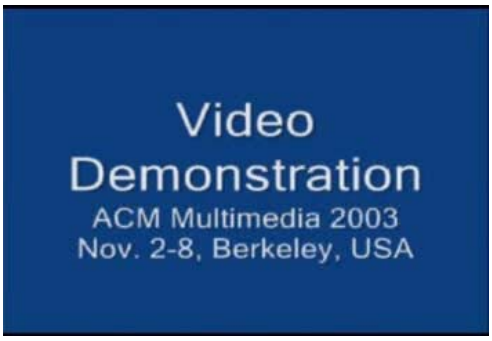
The **Photo2Video** system is a research prototype of Microsoft Research Asia in Beijing, China. It produces a video from a set of photos automatically.

Key technologies are

- the detection of regions of interest („focusses of attention“), using video content analysis algorithms
- the automatic generation of motion patterns, such as zooming in and panning, and appropriate trajectories
- the addition of incidental music and transitions between photos

Similar mechanisms can be used to “browse” a photo when the screen of the device is too small to show the entire image in an acceptable resolution.

Demo of the Photo2Video System



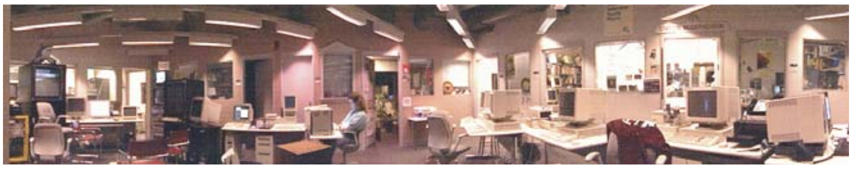
Video
Demonstration
ACM Multimedia 2003
Nov. 2-8, Berkeley, USA

8.8.4 Video to Panoramic Photo

When camera motion can be detected automatically, it is also possible to compute a panoramic image from a camera panning operation.

Example

Computation of panoramic images from a video that was recorded by a panning camera placed on a person‘s head (Steve Mann, MIT Media Lab, 1996)

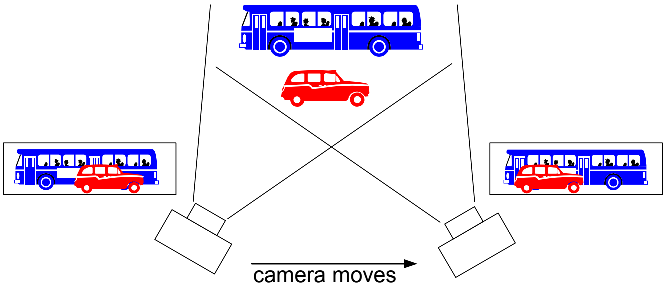


Panoramic Image Generation

We want to generate a panoramic image of our environment from a fixed position.

Parallax effect: If the camera moves objects at different distances appear shifted against each other.

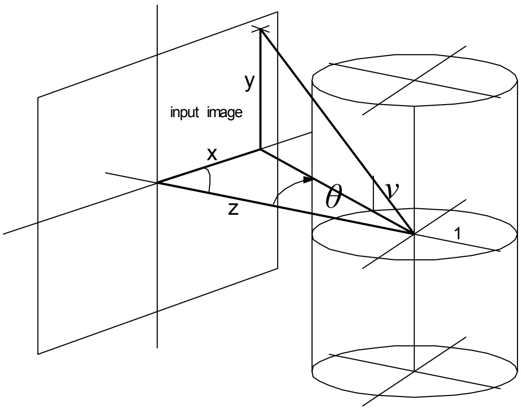
-> the camera should not change its position.



The easy approach: a **cylindrical panorama**. The camera only rotates around its vertical axis. Images are mapped onto a virtual cylinder around the camera position.

Cylindrical Panoramas (1)

Map the image to cylindrical coordinates.

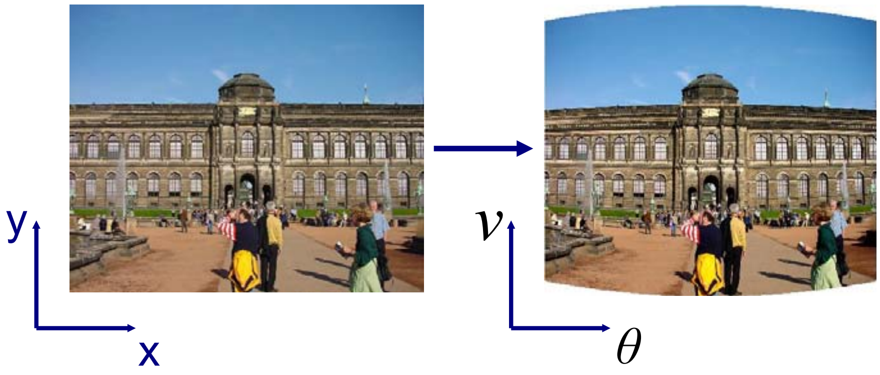


$$\theta = \arctan(x / z)$$
$$v = y / \sqrt{x^2 + z^2}$$

Note that the focal-length z has to be known for the coordinate transform.

Cylindrical Panoramas (2)

Image example



Cylindrical Panoramas (3)

Transformed images have to be aligned by matching the image content in the overlapping area:

Find the translation vector (t_x;t_y) by minimizing the matching error between images f and g:

$$Err(t_x, t_y) = \sum_{x,y} | f(x,y) - g(x+t_x, y+t_y) |^2$$

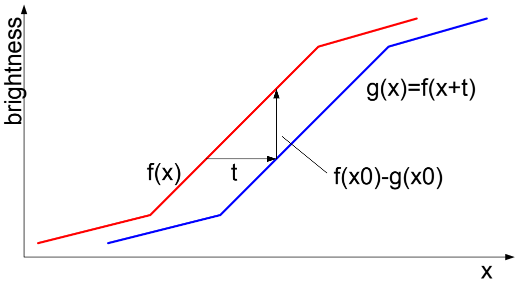
The minimum can be found by an exhaustive search over (t_x;t_y). However, computational complexity for image size *NxN* pixels and search range *MxM* pixels is:

$$O(N^2 M^2) \approx O(N^4)$$

We conclude that we need faster algorithms.

Cylindrical Panoramas (4)

Pel-Recursive Motion Estimation: For simplicity consider the one-dimensional case. Let g(x) be an exact copy of f(x), displaced by a constant t.



With the spatial derivative $\frac{\mathrm{d}f}{\mathrm{d}x}$ of f, we obtain t as:

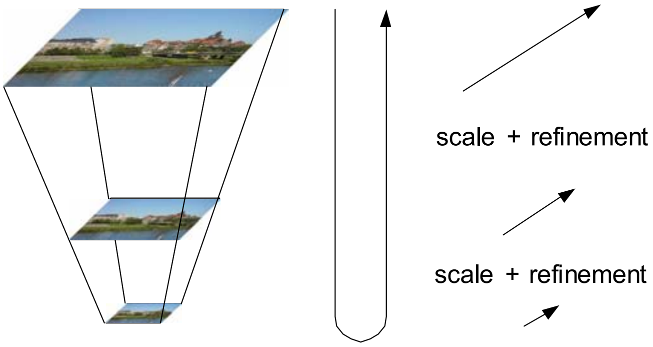
$$t = \frac{f(x_0) - g(x_0)}{\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)}$$

However, this only works for small displacements t.

Cylindrical Panoramas (5)

Hierarchical Motion Estimation - an Algorithm

- Scale the original image down by a constant factor to build a pyramid of images at different resolutions.
- Do motion estimation on the lowest resolution layer.
- Scale the obtained motion model upward to the next resolution level and refine the motion model.

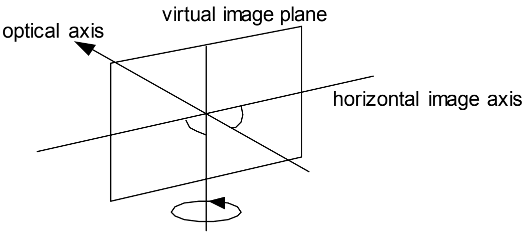


Cylindrical Panoramas (6)

Example result of a cylindrical panorama:

Disadvantages of cylindrical panoramas

- Distortion of straight lines to curved lines
- Focal length (zoom) has to be known or estimated
- Rotation axis has to be perpendicular to the optical axis and to the horizontal image axis:
  - Standing on a hill and looking down is not possible
  - Camera may not be rotated around horizontal axis (but mathematical extension to *spherical panoramas* is possible).



Full-Perspective Panoramas (1)

Assume that your environment is painted on a single, large plane (a glass plane).

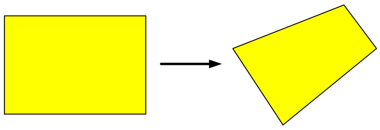
Describe the motion that the solid plane can perform in space.

2D: affine motion (translation, rotation, scaling)

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}}_{\text{rotate, scale}} + \underbrace{\begin{pmatrix} t_x \\ t_y \end{pmatrix}}_{\text{translate}}$$

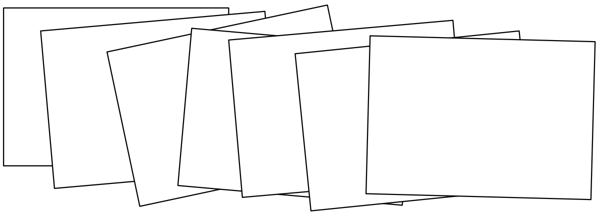
3D: perspective transformation (affine + two additional parameters for the perspective projection)

$$x' = \frac{a_{11}x + a_{12}y + t_x}{b_1 x + b_2 y + 1} \qquad y' = \frac{a_{21}x + a_{22}y + t_y}{b_1 x + b_2 y + 1}$$



Full-Perspective Panoramas (2)

Think of the camera taking images of the glass plane from different orientations. To align the images, the relative transformation between the images has to be estimated.



Exhaustive search for the parameters is not possible because the parameter space is very large (eight parameters).

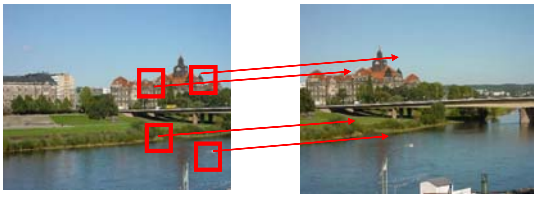
Possible approach for **parameter estimation:**

- Find an estimate with a lower-dimensional model (e.g., translatorial only to coarsely compensate motion)
- Determine an initial estimate for the perspective model (see next slide)
- Use the gradient-descent technique for fine alignment.

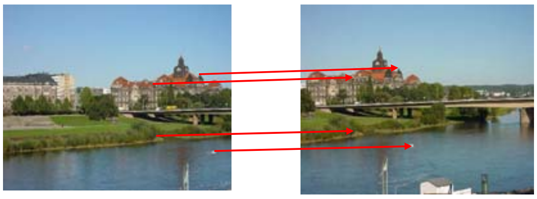
Full Perspective Panoramas (3)

Determine initial estimation of the perspective model:

1. Search for four blocks with characteristic features.



2. Even though we have compensated translatorial motion the features will not match perfectly. Refine matching positions of the blocks in the second image by block-matching. In general, the motion vectors will be different for the four blocks.



3. The four motion vectors (with two parameters each) can be used to determine the eight parameters of the perspective motion model (eight equations for eight unknowns).

Full-Perspective Panoramas (4)

Fine alignment

Start with the parameter vector estimated in last step:

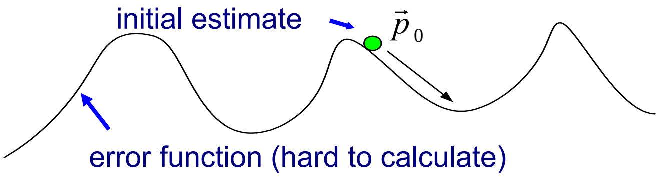
$$p_0 = (a_{11}; a_{12}; a_{21}; a_{22}; t_x; t_y; b_1; b_2)$$

2. Consider the matching-error depending on the transformation parameters:

$$Err(\vec{p}) = \sum_{x,y} | f(x,y) - g(x',y') |^2$$

3. Do a gradient descent search to find a better match.

$$\vec{p}_{i+1} = \vec{p}_i - \nabla Err(\vec{p}_i)$$



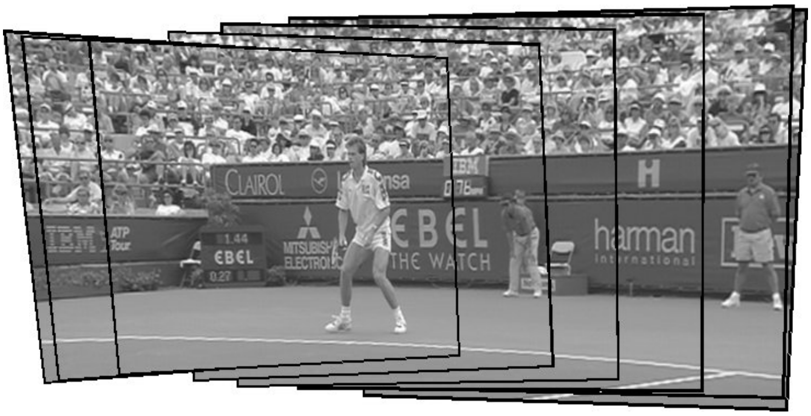
initial estimate

$\vec{p}_0$

error function (hard to calculate)

Full-Perspective Panoramas (5)

Image example

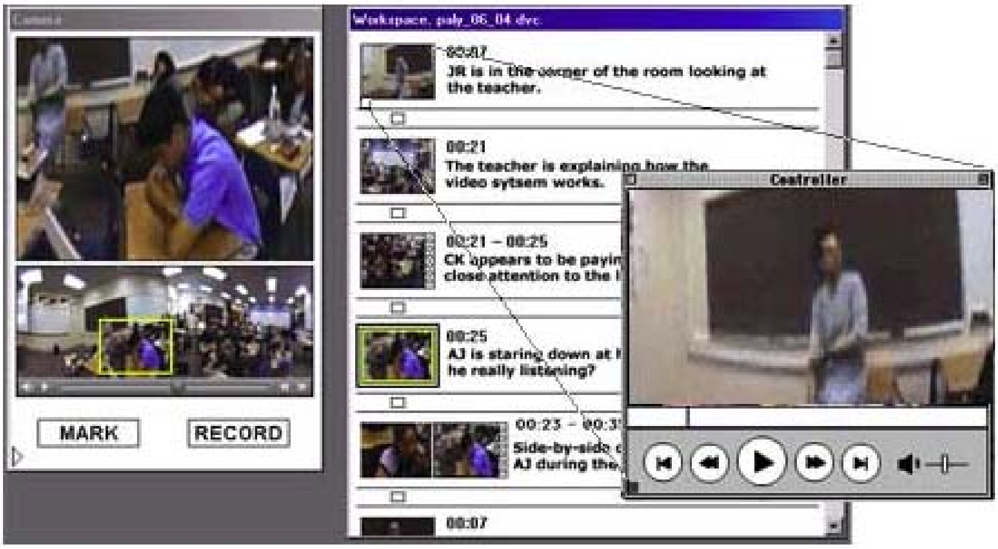


8.8.5 Repurposing of Panoramic Video: DIVER

- DIVER stands for **Digital Interactive Video Exploration and Reflection.**
- DIVER is a research prototype of Stanford University.
- It is a tool for authoring and sharing DIVEs. A DIVE is a unique perspective on a video record of human activity.
- DIVER separates the video recording process from the analysis and sharing process. It thus allows to re-purpose the video.

I thank Dr. Michael Mills and the entire DIVER team at Stanford University for letting me use their transparencies.

Example 1 of a DIVE: a Classroom



00:07
JR is in the corner of the room looking at the teacher.

00:21
The teacher is explaining how the video sytsem works.

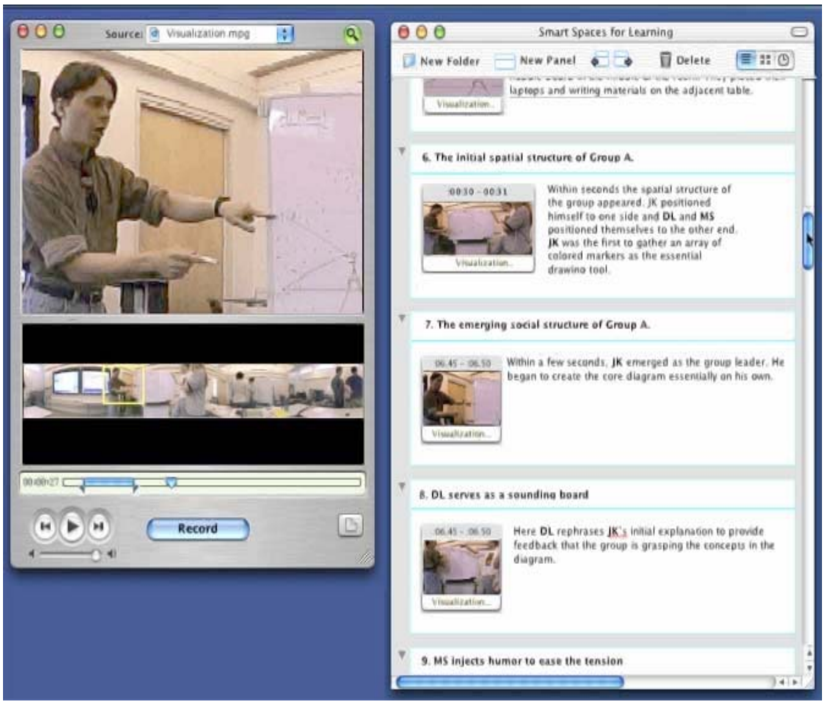
00:21 – 00:25
CK appears to be payin
close attention to the

00:25
AJ is staring down at
he really listening?

00:23 – 00:3
Side-by-side
AJ during the

00:07

Example 2 of a DIVE: a Meeting



laptops and writing materials on the adjacent table.

6. The initial spatial structure of Group A.

00:30 - 00:31
Within seconds the spatial structure of the group appeared. JK positioned himself to one side and **DL** and **MS** positioned themselves to the other end. **JK** was the first to gather an array of colored markers as the essential drawing tool.

7. The emerging social structure of Group A.

06:45 - 06:50
Within a few seconds, **JK** emerged as the group leader. He began to create the core diagram essentially on his own.

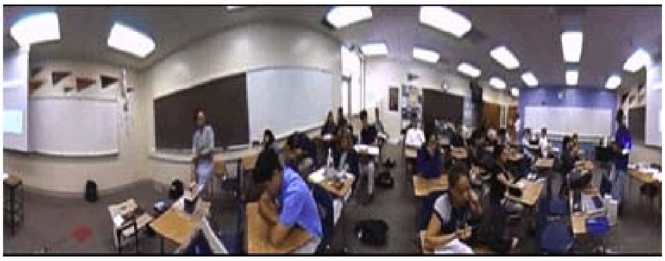
8. DL serves as a sounding board

06:45 - 06:50
Here **DL** rephrases JK's initial explanation to provide feedback that the group is grasping the concepts in the diagram.

9. MS injects humor to ease the tension

Panoramic Video in DIVER

- DIVER supports panoramic video.
- 360 degree views allow video recording of interesting activities without knowing the exact research goal. This is a good basis for re-purposing.
- Panoramic video at full resolution and a high frame rate is a major technical challenge!



What problem is DIVER solving?

Video is an important basis for research in human activities.

Example: classroom teaching

- What does the novice teacher notice?
- What does the expert teacher notice?
- What do different researchers notice?

Enabling Non-Selective Capture

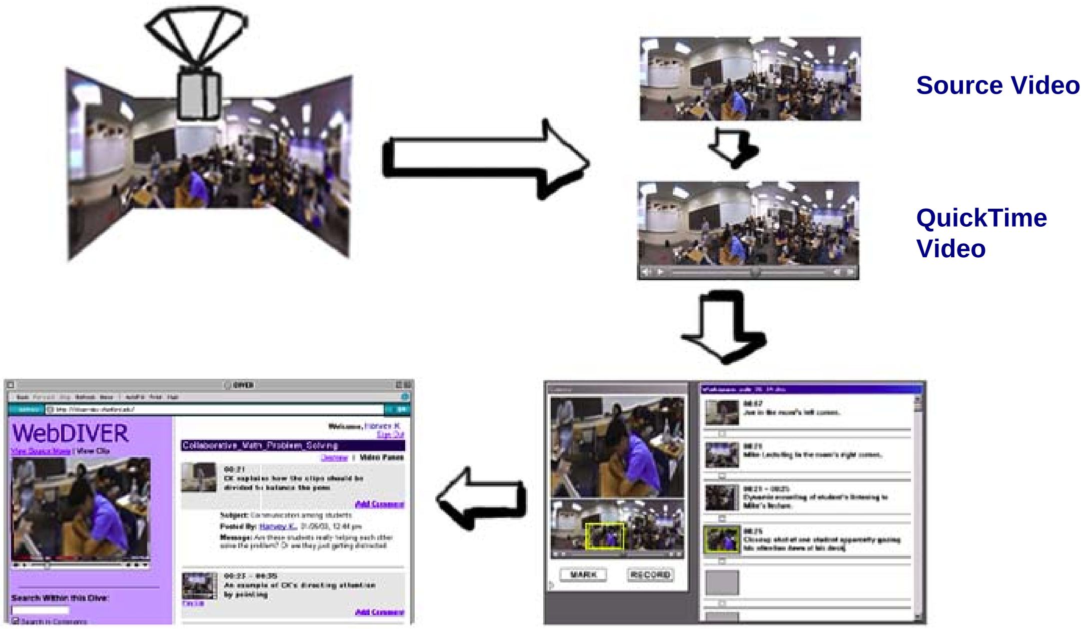
Selective capture is a major drawback in traditional video research.

A single camera and microphone often miss events occurring outside the field of view.

Main Goals of DIVER

- Support panoramic video to overcome the problem of selective capture.
- Develop user interfaces to support exploring the same event from different perspectives. Find user interfaces for panoramic video that are easy to understand and navigate.
- Develop “digital video collaboratories” – communities for annotating, sharing and discussing perspectives on digital video records

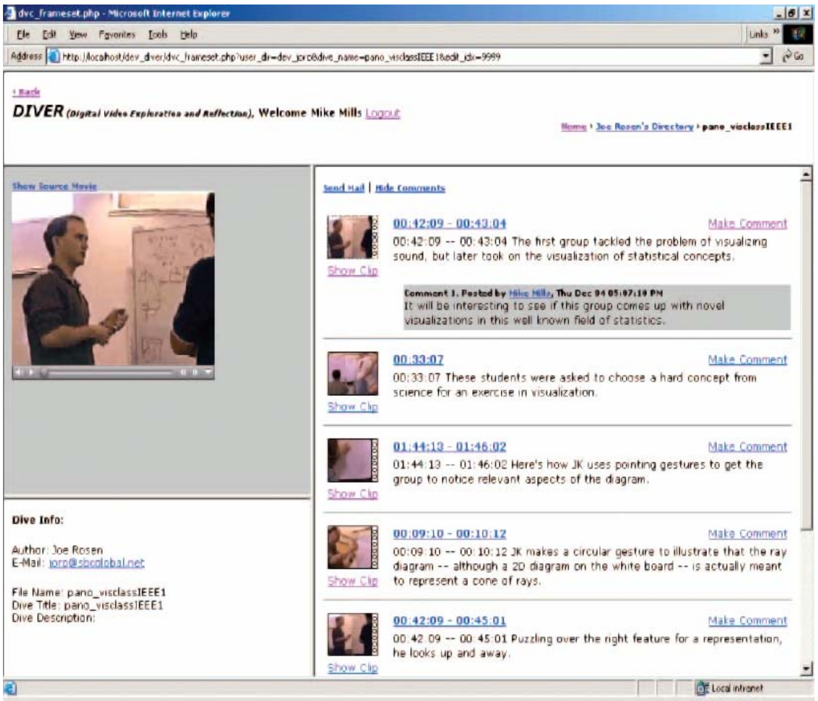
DIVER System Architecture



Source Video

QuickTime Video

Web DIVER (“Video Collaboratory“)


dvc_frameset.php - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Links
Address
Go
Local intranet

http://localhost/dev_diver/dvc_frameset.php?user_dir=dev_jorc&dive_name=pano_visclassIEEE1&edit_idx=9999

‹ Back

DIVER *(Digital Video Exploration and Reflection)*, Welcome Mike Mills Logout

Home › Joe Rosen's Directory › pano_visclassIEEE1

Show Source Movie

Send Mail | Hide Comments

00:42:09 - 00:43:04 Make Comment

00:42:09 -- 00:43:04 The first group tackled the problem of visualizing sound, but later took on the visualization of statistical concepts.

Show Clip

Comment 1. Posted by Mike Mills, **Thu Dec 04 05:07:10 PM**
It will be interesting to see if this group comes up with novel visualizations in this well known field of statistics.

00:33:07 Make Comment

00:33:07 These students were asked to choose a hard concept from science for an exercise in visualization.

Show Clip

01:44:13 - 01:46:02 Make Comment

01:44:13 -- 01:46:02 Here's how JK uses pointing gestures to get the group to notice relevant aspects of the diagram.

Show Clip

00:09:10 - 00:10:12 Make Comment

00:09:10 -- 00:10:12 JK makes a circular gesture to illustrate that the ray diagram -- although a 2D diagram on the white board -- is actually meant to represent a cone of rays.

Show Clip

00:42:09 - 00:45:01 Make Comment

00:42:09 -- 00:45:01 Puzzling over the right feature for a representation, he looks up and away.

Show Clip

Dive Info:

Author: Joe Rosen
E-Mail: jorc@sbcglobal.net

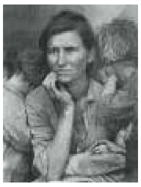
File Name: pano_visclassIEEE1
Dive Title: pano_visclassIEEE1
Dive Description:

Different DIVER Source Formats

High-resolution panoramic video

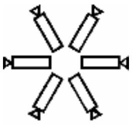
Low-resolution panoramic video

Traditional 4:3 aspect ratio video



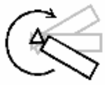
Diving on a still image

Recording Panoramic Video Is Not Easy!



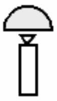
Multiple cameras, each looking out into the scene

- Individual images cannot be integrated seamlessly, problems with objects at short range
- Time-consuming complex image processing that blurs the images.



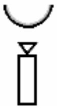
A single camera looking in different directions at different instants of time

- Images of moving objects are distorted, fragmented, or missing
- Real panoramic video is not possible.



A single camera looking out through a single panoramic lens

- Images of objects that are close to the camera are distorted
- planar projections from curved mirrors or fish-eye lenses do not evenly or fully cover the pixel array



A single camera looking off a single non-planar mirror

- Images are captured with low resolution
- Images have low optical quality
- Scene must be very brightly illuminated or static

Fullview‘s Panoramic Camera (1)

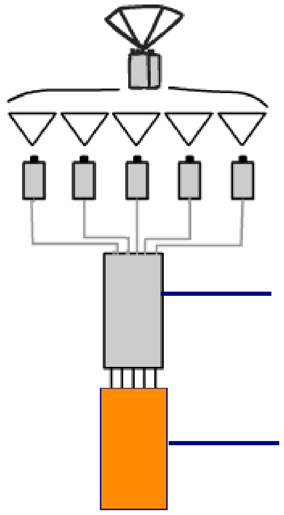
Multiple cameras, each looking into a mirror. Mirrors are arranged in a pyramid.

- All cameras have exactly the same point of view.
- Individual images can be integrated seamlessly.
- Image resolution is high.

But: Stitching, de-warping, color correction have to be done in real-time. And the data rate is very high.

Fullview‘s Panoramic Camera (2)

Five high-resolution FC-1005 cameras views five overlapping scene regions reflected off planar mirrors. The cameras output five S-video signals to the Camera Control Unit (CCU) which is controlled by a PC.



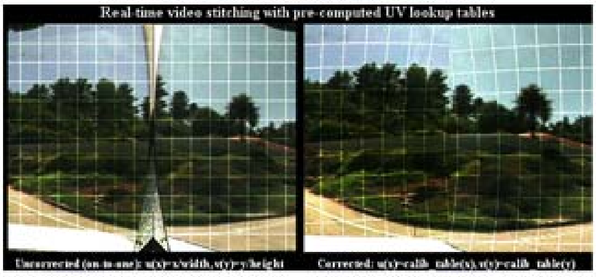
CCU controls the cameras (white balance, shutter, gain)

PC with 5 frame grabbers controls CCU, serves as a compute engine for stitching, de-warping

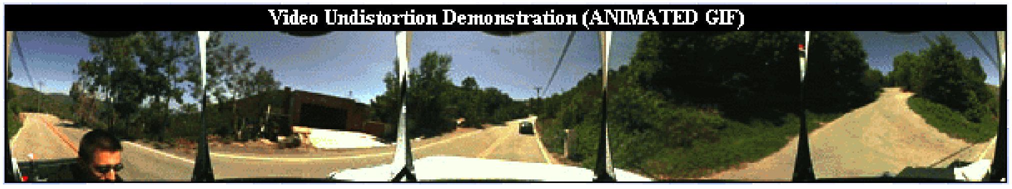
Technical Challenges

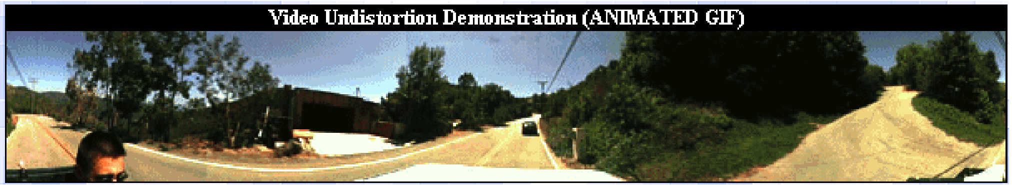
- Stitching and de-warping in real-time
- User interface to view and navigate panoramic video
- Data transfer to disk at full resolution and full frame rate (uncompressed)
- Multi-channel audio

Stitching and De-Warping

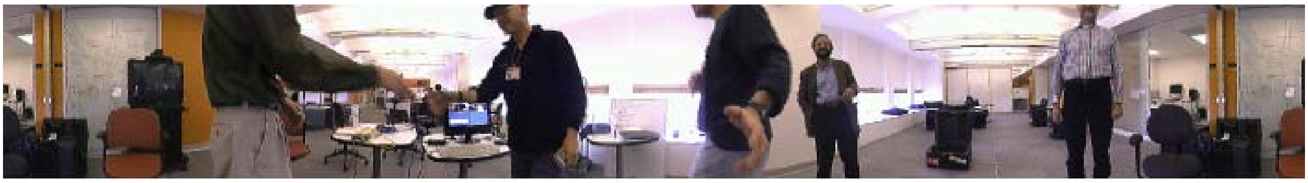

Real-time video stitching with pre-computed UV lookup tables
Uncorrected (on-to-one): u(x)=x/width, v(y)=y/height
Corrected: u(x)=calib_table(x), v(y)=calib_table(y)

- Find texture similarities or edge similarities in the overlap
- Parameterize a matrix to transform the images (example courtesy of USC)
- Computationally very expensive!





Example of a Fullview Panoramic Image



Current Problems

- Panoramic video cannot be written to disk uncompressed, at full rate and full resolution over a PCI bus (half-resolution at 12 fps require 3 PCI buses!)
- No audio yet (10-microphone directional audio planned)
- Moving the camera to a different room requires lengthy re-calibration of the white balance for the five cameras (but not of the frame stitching parameters)

User Interfaces for Panoramic Video

- The human brain is trained to understand the visual input from the eyes as coming from the normal angle of vision (150 degrees?)
- It is very difficult to understand the 3D relationship between objects in a flat window on the screen. Example: Who has eye contact with whom?
- **Open problem:** What would be a good user interface to view and navigate a panoramic video on a flat screen?

DIVER Summary

- The use of DIVER with panoramic videos enables “virtual videography” and overcomes many current problems with selective video capture.
- “Path movies” that dynamically crop focus within digital video can provide a crucial new cultural medium for expressing one’s perspective.
- A first release of the DIVER system, including Web-Diver, is fully operational.
- First experiments with educational researchers were conducted in Winter 2002/03.
- Full resolution, full frame rate panoramic video recording and multi-track directional audio recording planned.