Cross-modal mappings and audiovisual feedback

Nick Collins

University of Sussex

Summary

- Experiments in audiovisual combination and data flow
- Challenges of audiovisual mapping

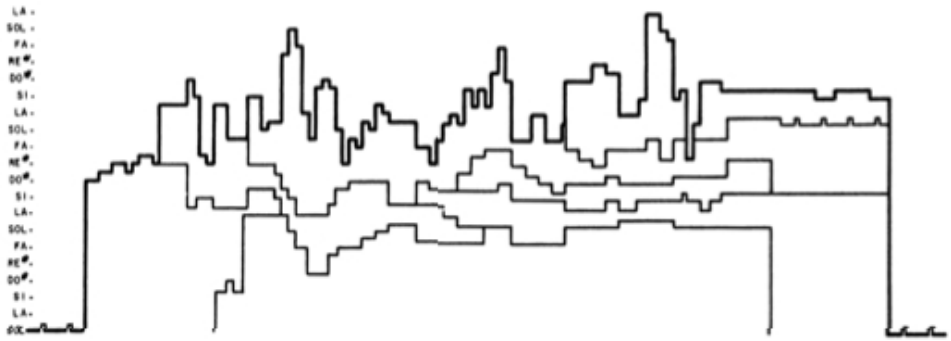
Theoretical backdrop

- Film music vs music film
- 1-1 mappings, mickey-mousing, synchresis
- Meanings in parallel, audiovisual objects

Problems with mapping

- Maps between domains are usually not homomorphisms; not structure-preserving
- Sonification/visualisation from data, audio driving visuals or vice versa; all examples of potentially incompatible spaces

VILLA-LOBOS: NEW YORK SKY LINE MELODY - GRÁFICO DERIVADO DA VERSÃO DE 1957. (C. KATER, 1982)



Psychological backdrop

- Synaesthesia is personal, not universal
- Audiovisual integration, attentional resources
- Coping with multiple objects in counterpoint

Some successful close audiovisual integrations?

- Coldcut/Hex *Timber* (1997)
- Autechre/Alex Rutterford *Gantz Graf* (2002)
- chdh's audiovisual objects

Now…

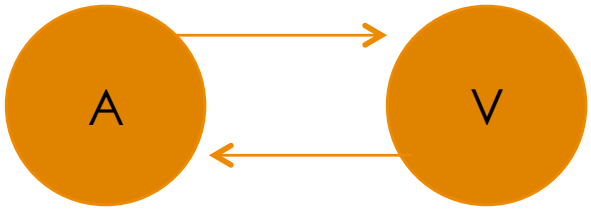
- Some unsuccessful relations?
- Case studies from experiments

Case 1: PhotoNoise iPad app

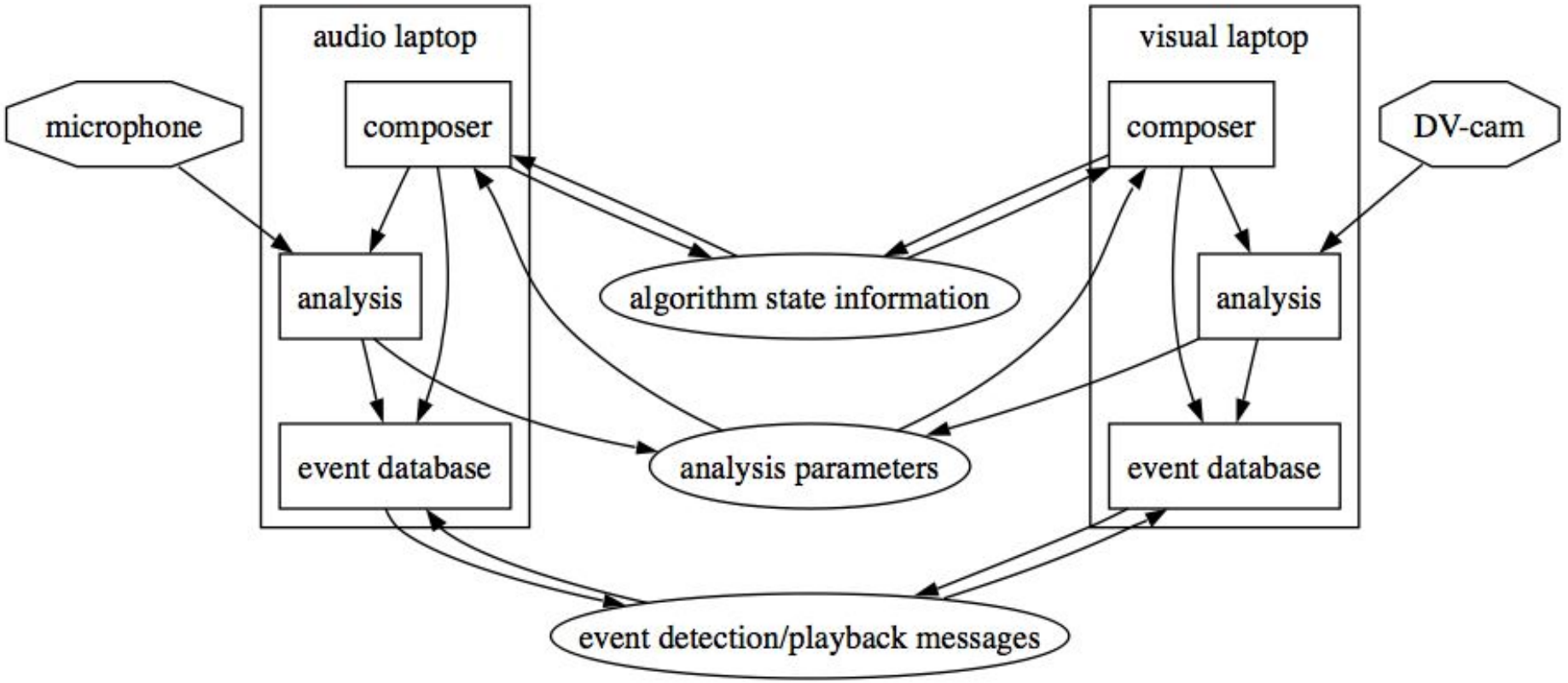
- Image -> sound via pixel data
- Multi-touch -> RGB + intensity -> nonlinear oscillators
- Quirky, noisy, more of an esoteric instrument than deep in itself

Audiovisual feedback

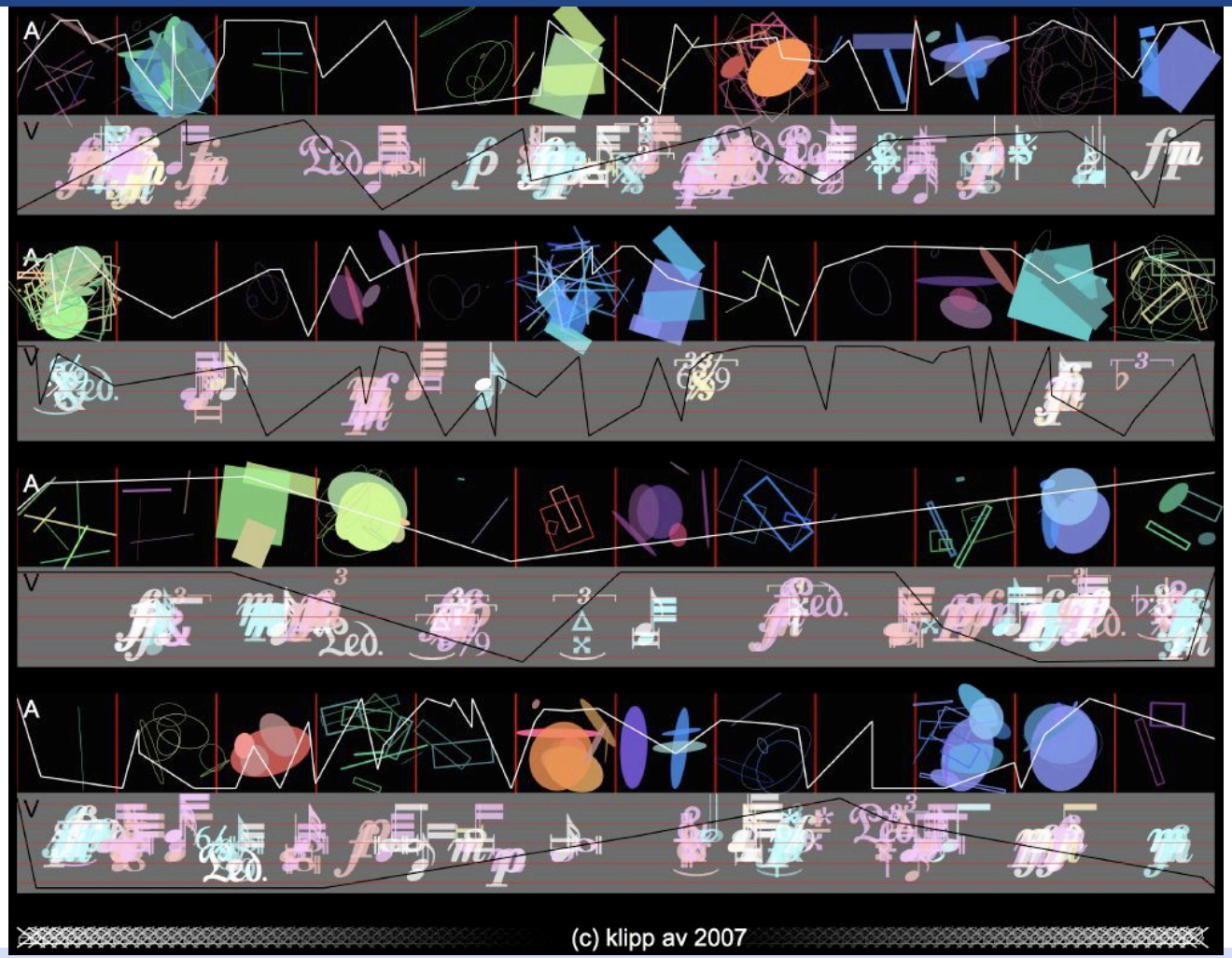
- Round a processing chain incorporating both audio and video material and its analysis
- E.g., audio input -> audio analysis -> drive generative graphics -> video analysis -> drive audio synthesis -> audio input …
- Processing delays



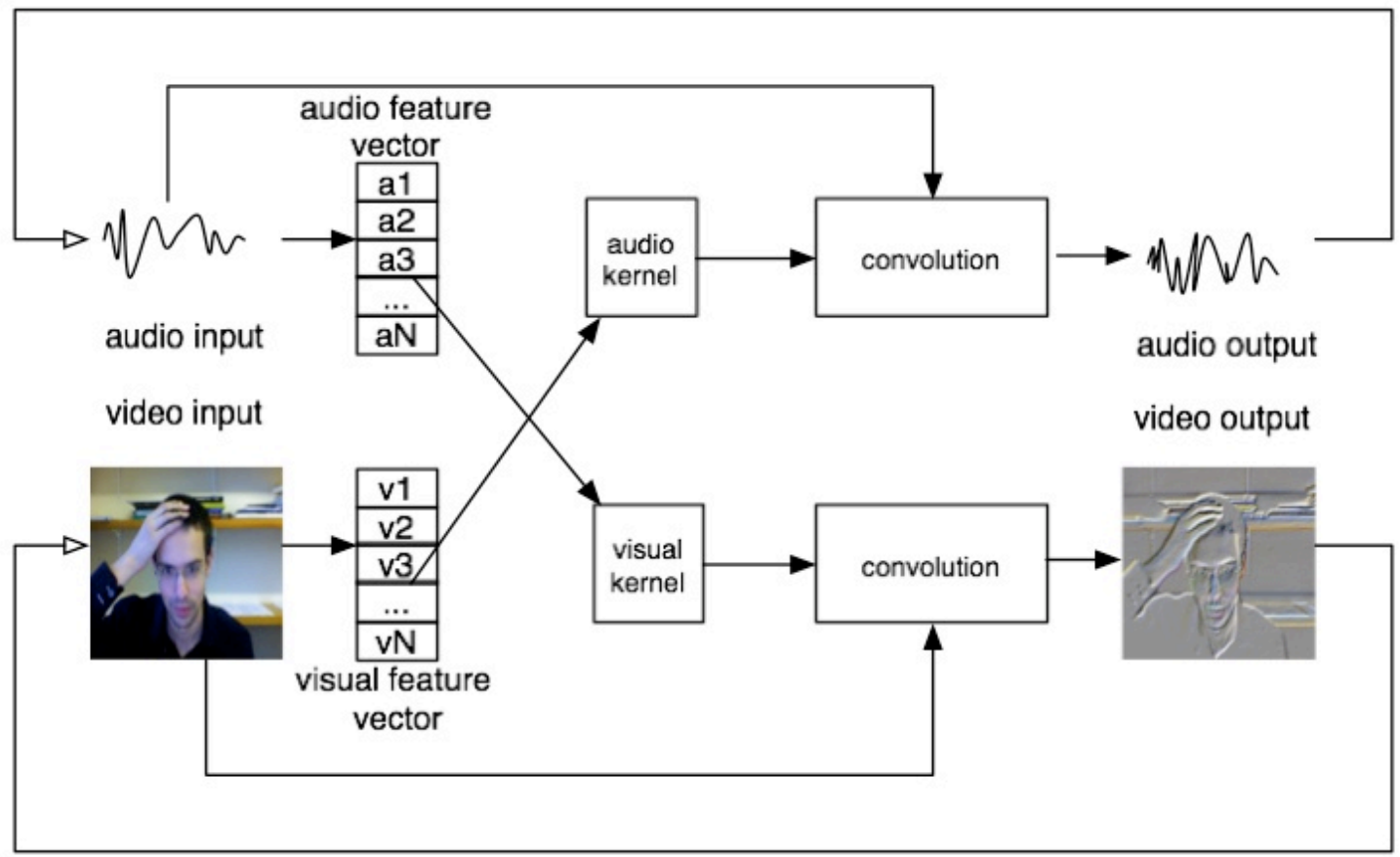
Case 2: klipp av (2002-2009)



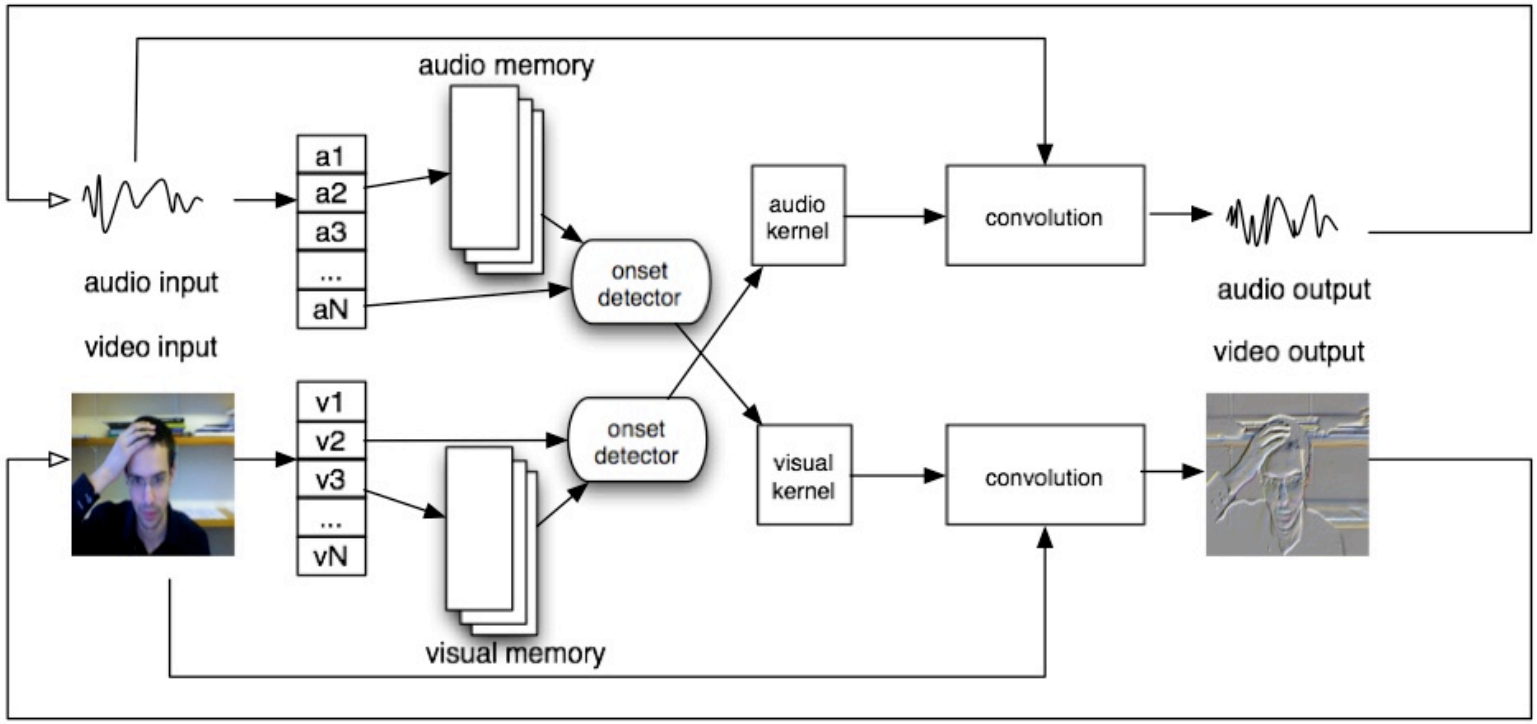
Audiovisual performance score



Case 3: iterative cross-convolution

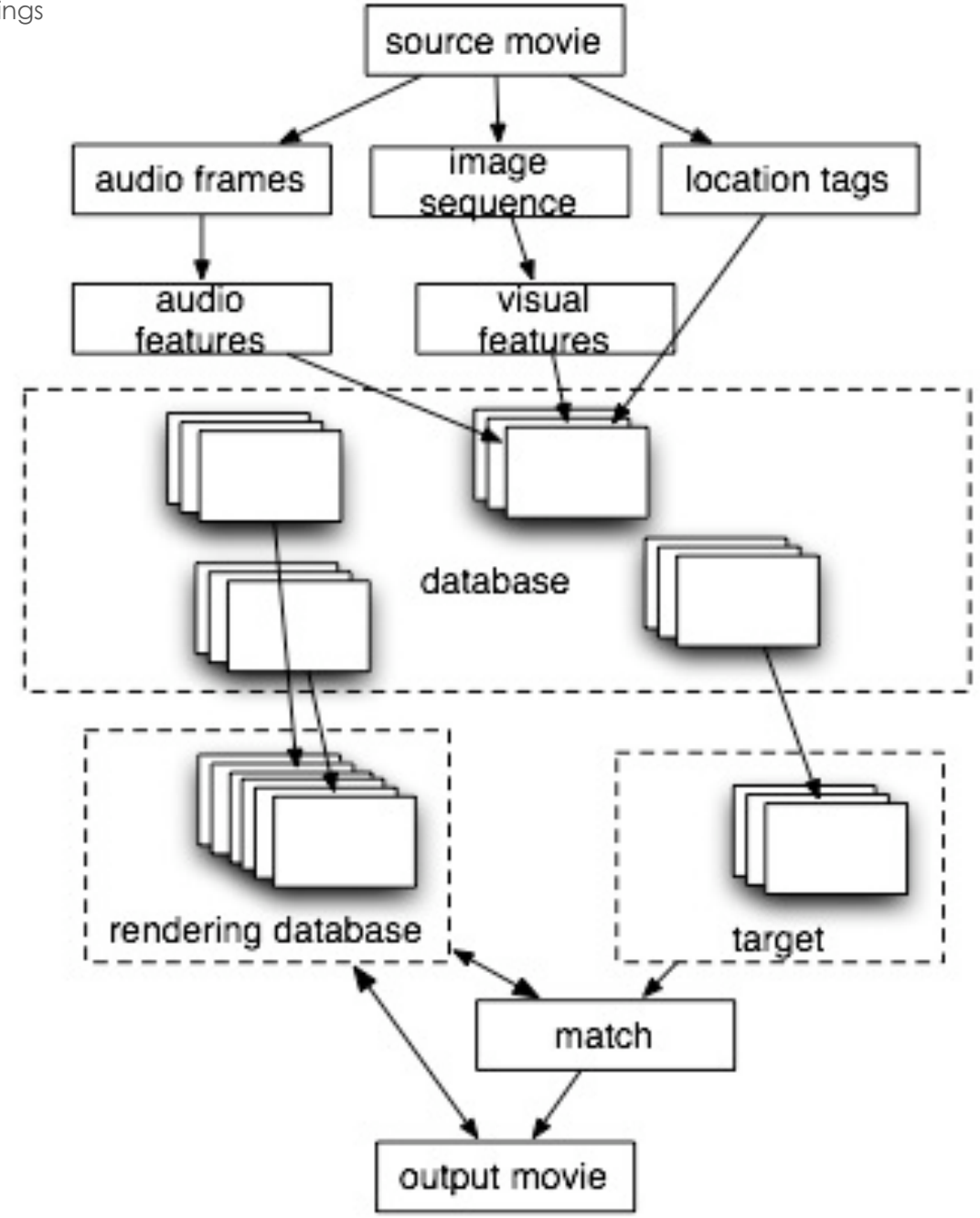


With onset detection



Case 4: audiovisual concatenative synthesis

- Concatenative synthesis = feature-matched substitution of an input from a database
- Usually audio or visual alone.
- Has been explored where video is brought along for free (Grierson and Casey 2007, Scrambled Hackz et al.)
- Can do both at once = audio and video feature data muddle


source movie
audio frames
image sequence
location tags
audio features
visual features
database
rendering database
target
match
output movie

Case 5: auto aligned video

- Take an existing music work
- Extract audio features -> drive frame by frame rendering of generative graphics
- End up with close correlation
- Tool chain: SCMIR->Processing->ffmpeg

What is coming out of this?

- 1-1 mapping might be simple, but it is perceptually strong!
- Need for audiovisual objects as coherent entities
- Details of many mappings lost (can be other reasons to keep on regardless, e.g. improvisation in complexity)

Conclusions

- Considered some of the challenges of correlating audio and visual
- Discussed some projects of varying effectiveness
- Details often lost when mapping between modalities

Think you for lastening

http://www.sussex.ac.uk/Users/nc81/researchav.html

Alexander, A., and Collins, N. (2007) Live Audiovisuals. Book chapter in N.Collins and J. d'Escrivan (Eds.) (2007) The Cambridge Companion to Electronic Music

(2007) "Audiovisual Concatenative Synthesis". Proceedings of ICMC2007, International Computer Music Conference, Copenhagen.

(2006) Nick Collins and Fredrik Olofsson. "klipp av: Live Algorithmic Splicing and Audiovisual Event Capture", Computer Music Journal 30(2): pp 8-18

http://www.sussex.ac.uk/Users/nc81/iphone.html#PhotoNoise