Temporal PsychoVisual Modulation (TPVM)

A breakthrough in information display technology

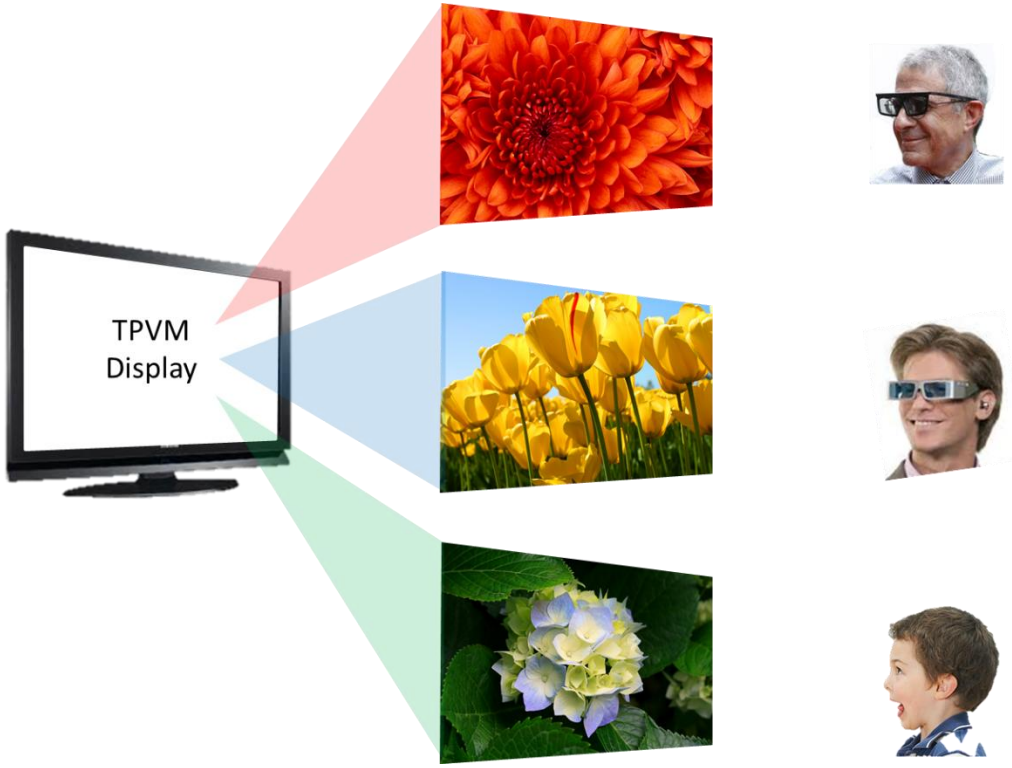
Xiaolin Wu
Dept. of Electrical & Computer Engineering
McMaster University
Hamilton, Ontario, Canada

featured in MIT Technology Review:
http://www.technologyreview.com/printer_friendly_blog.aspx?id=27490
Demos: http://www.quasion-inc.com/tpvm

One Exhibition per Display

One thing in common from tribesmen's rock carving to today's 3D displays

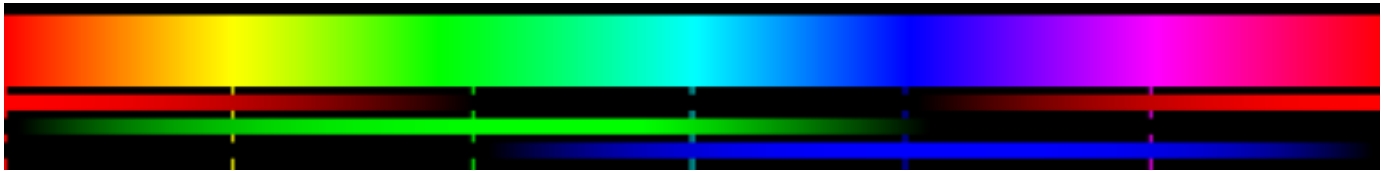
Concurrent Multiple Exhibitions on a Single TPVM Display



A new paradigm of information display: TPVM differs fundamentally, in principle, design, user experience and cost effectiveness, from MIT's head-mounted display technology and from the dual-view technology (TI and Sony's patents).

How Could One-for-Many Display Possible (眼见为实)?

How much can we see?



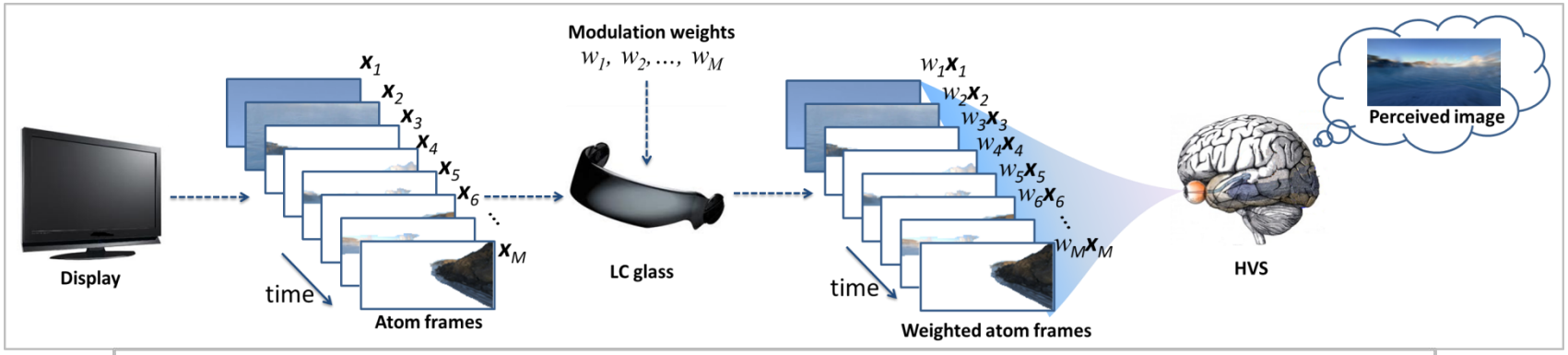
Capacity of human visual cognition – limited resolution of human visual system (HVS) in spatial, spectral, and temporal domains.

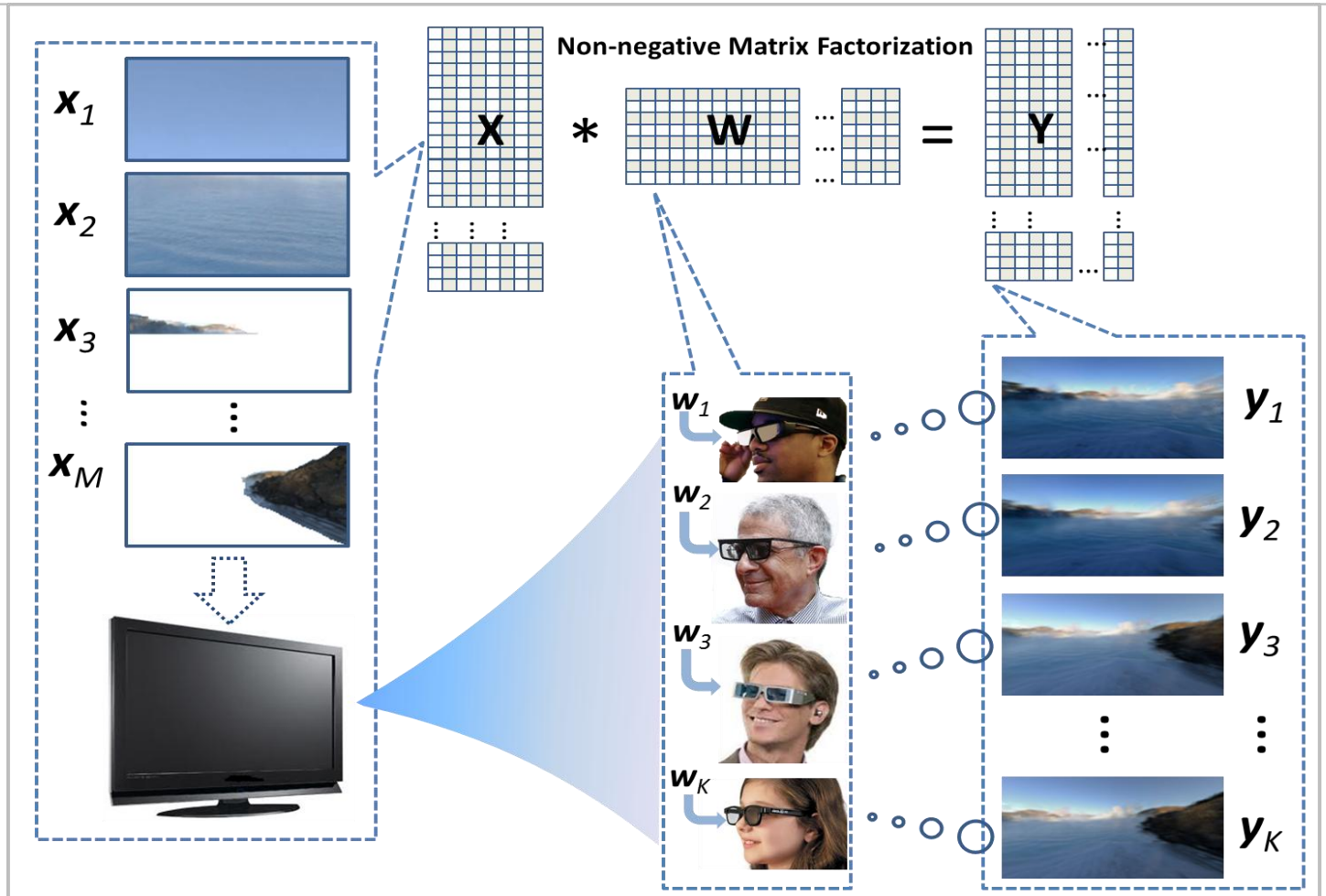
Can we see everything in visible spectrum?

Critical flicker frequency is about 60Hz for most people, while the modern displays can run at 120Hz and beyond.

A high-speed display can emit in visible spectrum a far greater amount of information than one person can possibly resolve.

Scientific Principle Of TPVM





Hybrid Optoelectronic-Psychovisual Image Formation

- A single high speed display broadcasts a set of basis frames.
- LC glasses linearly weigh the basis frames.
- Different views are formed by the human visual systems that fuse differently weighed basis frames.
- Computational image formation via an ingenious interplay of cognitive science, optoelectronics, and signal processing.

Problem Formulation

f_d --- display refresh rate

f_c --- flicker fusion rate, assuming $f_d = Mf_c$

$\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ --- K concurrent views

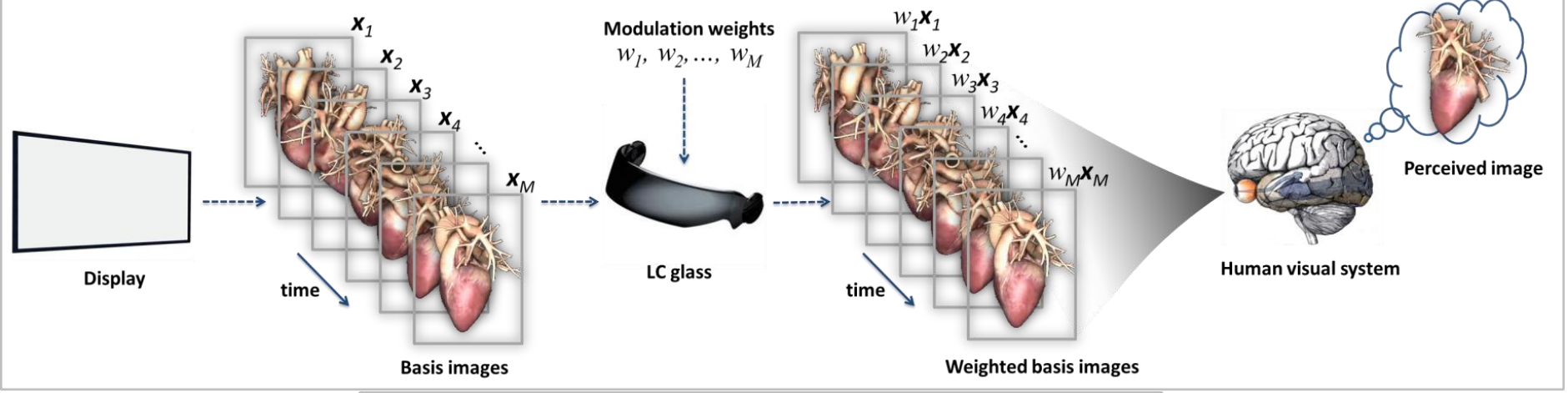
The K target images are represented by Y with its k-th column being image $\mathbf{y}_k$.

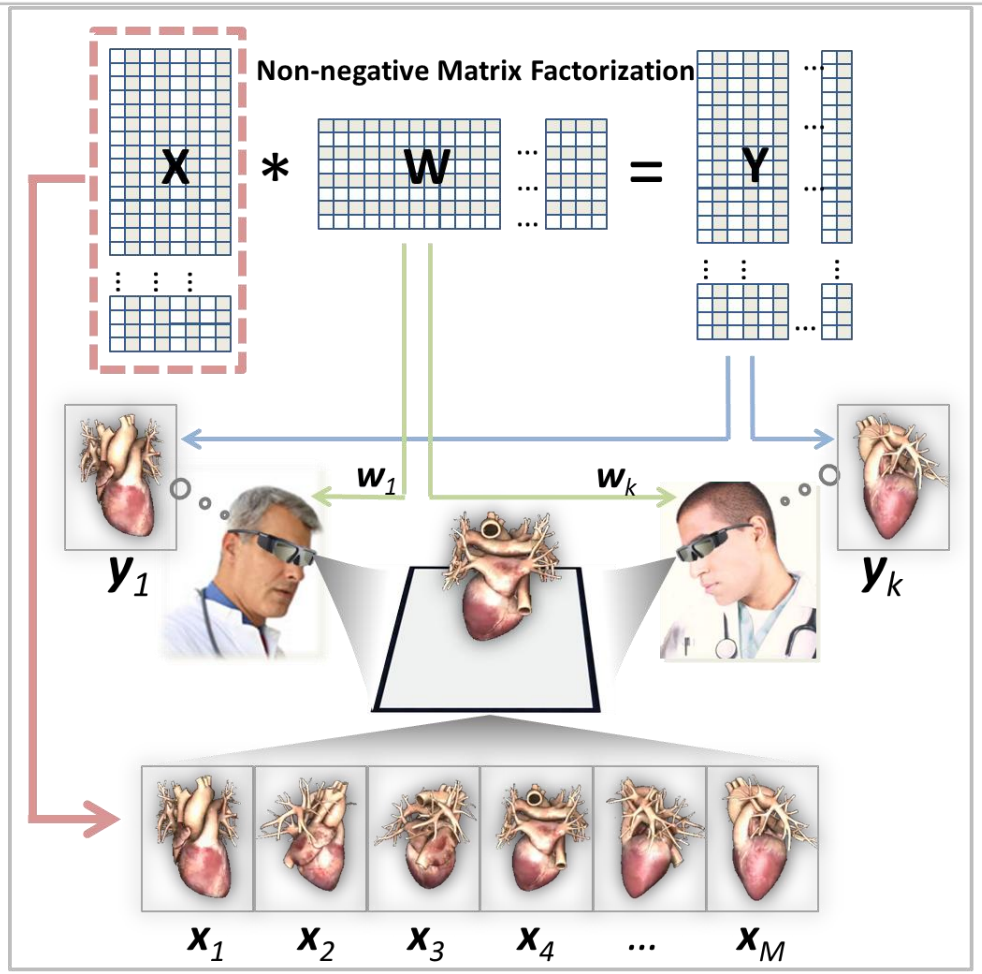
TPVM can be formulated as a problem of signal decomposition

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{W}$$

$\boldsymbol{X}$ --- $N \times M$ matrix with columns being basis M images $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M$

$\boldsymbol{W}$ --- $M \times K$ matrix of modulation weights.


Display
time
Basis images
Modulation weights
$w_1, w_2, \ldots, w_M$
LC glass
time
Weighted basis images
Human visual system
Perceived image


Non-negative Matrix Factorization
X
W
Y
w_1
w_k
y_1
y_k

TPVM via Non-negative Matrix Factorization

Light energy cannot be negative and active LC glasses cannot apply negative weights

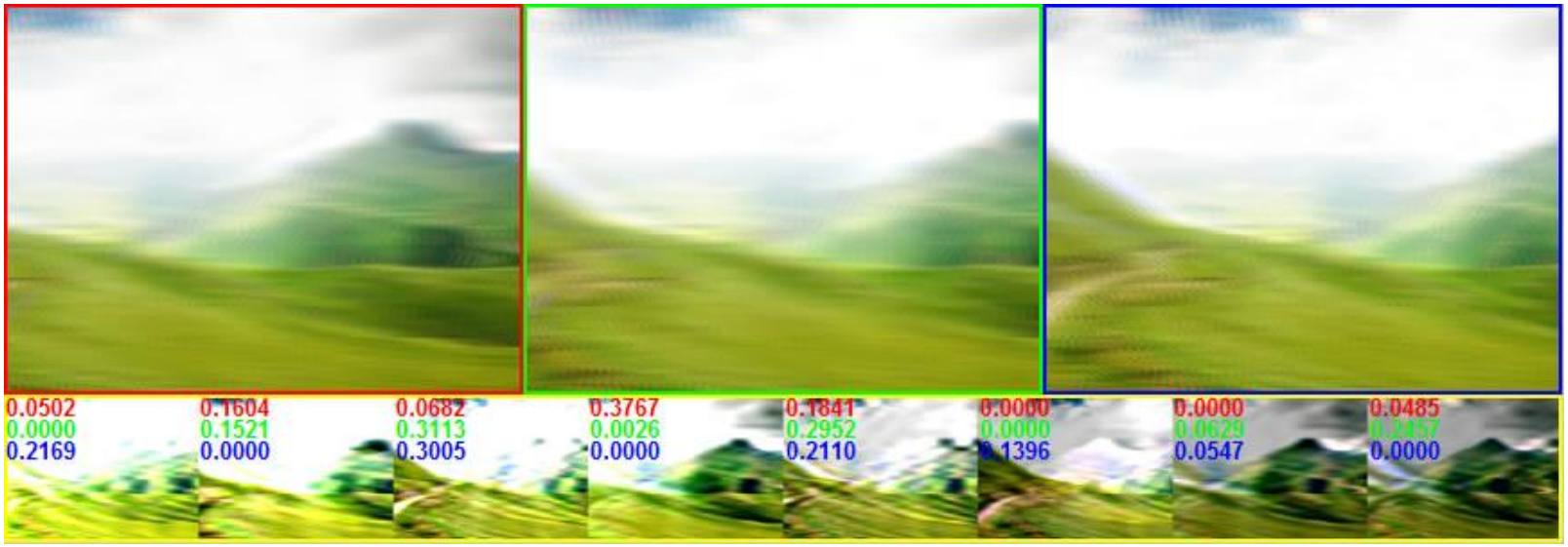
The signal decomposition $Y = XW$ has to be a non-negative matrix factorization (NMF).

The TPVM display system solves the following optimization problem

$$\min_{X,W} \left\| sY - XW \right\|_F \quad s.t. 0 \le X \le 1, 0 \le W \le 1$$

The scaling factor s, $0 \le s \le M$, ensures each user receives sufficient energy level (brightness).

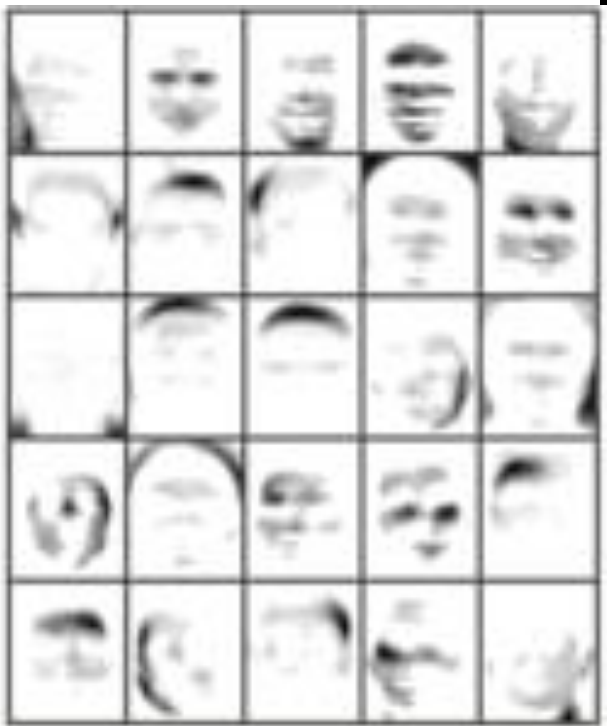
An Example


0.0502
0.0000
0.2169
0.1604
0.1521
0.0000
0.0682
0.3113
0.3005
0.3767
0.0026
0.0000
0.1841
0.2952
0.2110
0.0000
0.0000
0.1396
0.0000
0.0629
0.0547
0.0485
0.0000

Neurological and Cognitive Backups for TPVM

Visual cognition is based on a brain scene representation in terms of parts [Biederman ‘87; Ullman ‘96], with psychological and physiological evidence [Palmer ‘77; Wachsmuth et al. ‘94; Logothetis and Sheinberg ‘96].

Basis images flash very quickly (8.3ms for 120 Hz displays), each likely contains a meaningful part of a primed scene. In such cases, viewers respond positively to transient visual stimuli [Thorpe et al. ‘96].



Glass Free Rendering

What happens to those who watch the same screen not through any modulated viewing devices?

The image $\mathbf{y}_0 = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_M$ (*normal view)* results, rendered by HVS of all basis images when they are displayed in rapid succession and perceived unattenuated.

The normal view should be semantically meaningful and visually pleasing in addition to other intended glass-aided views (*shale views*).

This functionality, if attainable, is much desired or even necessary in many VR applications

K+1 Viewing (I)

Given the set of K shale views in Y and a normal view $\mathbf{y}_0$, compute the M generation basis images $\tilde{X}$ and K weighting vectors to satisfy

$$Y = \tilde{X}\tilde{W}; \mathbf{y}_0 = \tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_2 + \ldots + \tilde{\mathbf{x}}_M$$

First, solve the original NMF problem for TPVM and obtain X and W such that

$$Y = XW$$

The NMF solution is not unique in general, because

$$XW = XDD^{-1}W$$

for any invertible matrix D.

K+1 Viewing (II)

From $XW = XDD^{-1}W$, we can let

$$\tilde{X} = XD, \tilde{W} = D^{-1}W$$

The normal view $\mathbf{y}_0$ can be approximated as a linear combination of the bases

$$\mathbf{y}_0 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots c_M\mathbf{x}_M$$

Let

$$D = diag(c_1, c_2, \ldots, c_M)$$

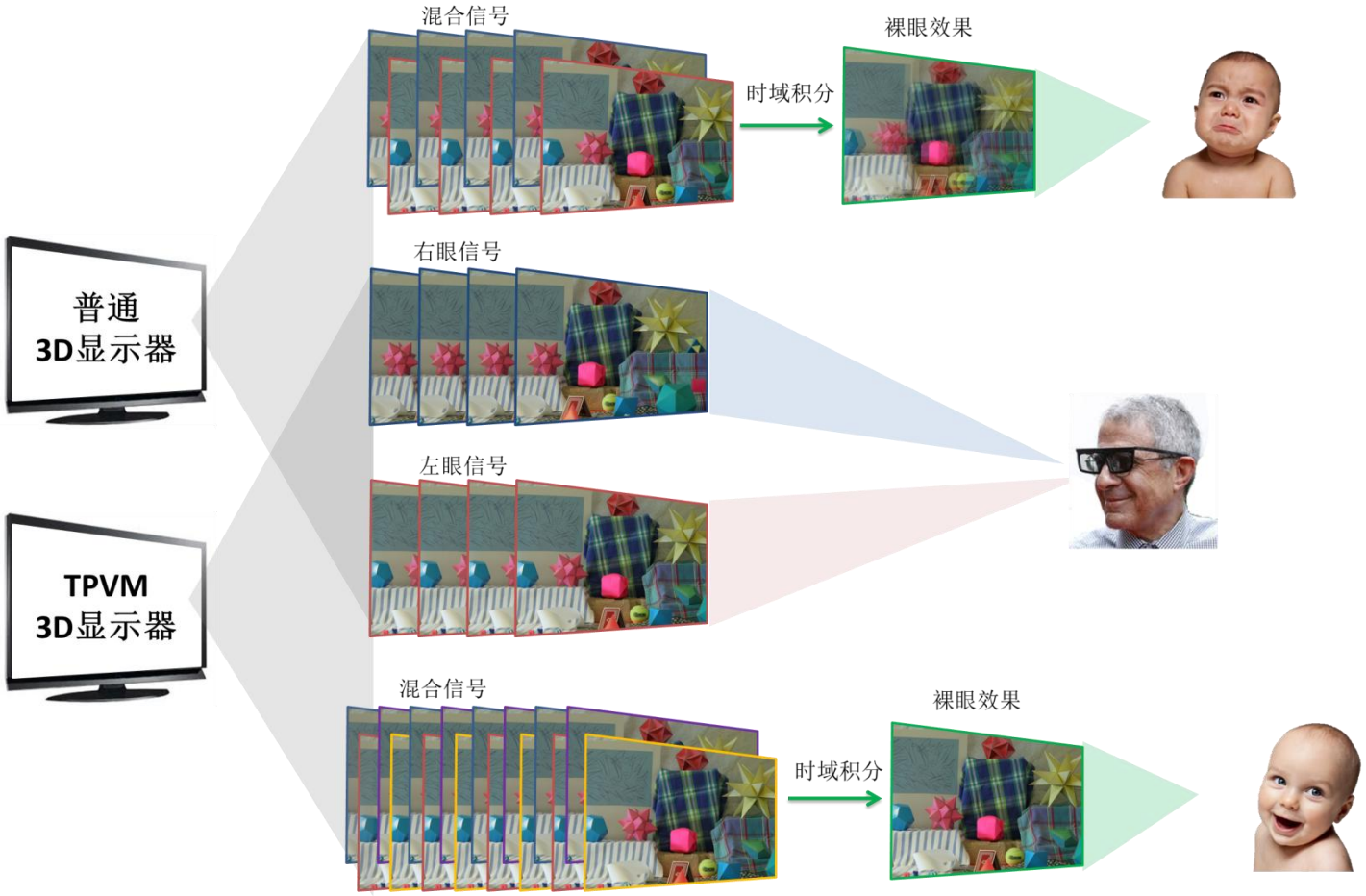
where $c_1, c_2, \ldots, c_M$ are the solution of the constrained least square problem

$$\min_{\mathbf{c}} \left\| \mathbf{y}_0 - W\mathbf{c} \right\|_2^2 \quad s.t. \ \mathbf{c} \geq 0$$

The constraint $\mathbf{c} \geq 0$ ensures the nonnegativity of $\tilde{X}$ and $\tilde{W}$.

2D and 3D Concurrent Viewing (I)

A special case: $K=2$ shale views are the left-eye and right-eye signals of a 3D scene; the glass-free normal view is a conventional 2D presentation of the same scene. Immediately, a TPVM solution to open problem of 2D and 3D concurrent viewing.



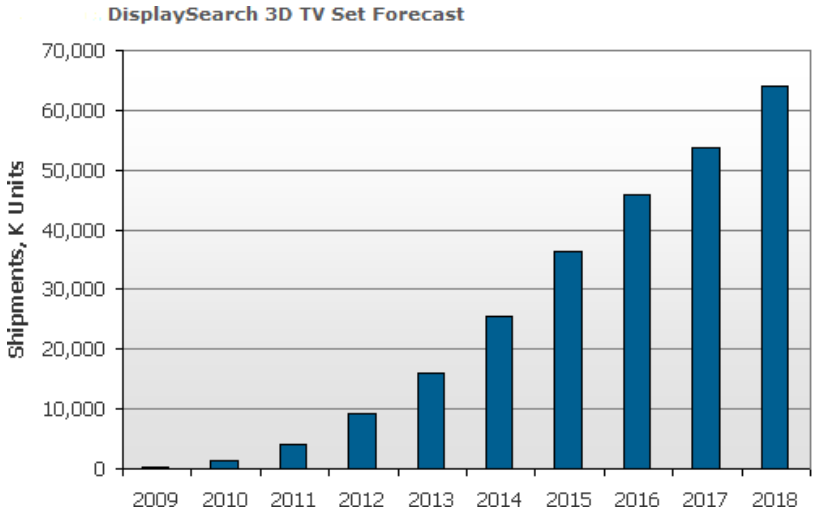
3D Impetus

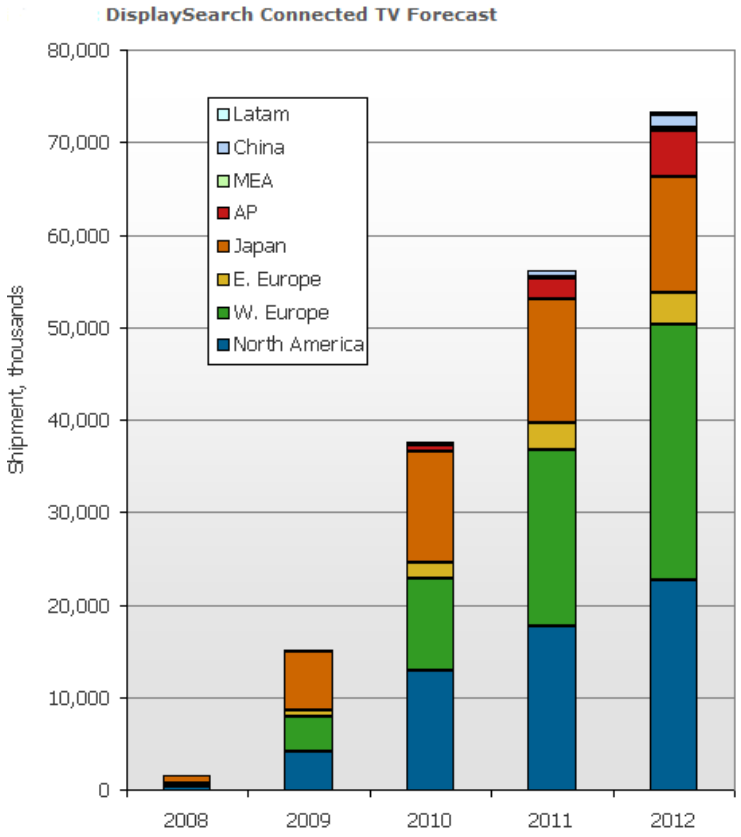
Active shutter glasses type (current industrial standard), refresh rate ≥120 Hz

- (Reuters) - The global 3D television market will grow more than 5-fold to account for 11 percent of flat-screen TV sales in 2011, as prices fall sharply and 3D becoming a common add-on.
- 3D TV shipments will rise to 23.4 million units in 2011, from last year's 4.2 million units, gaining further to 159 million units in 2015. By that time, more than half of global flat-panel shipments.

The Trend of 3D TV

Research firm, DisplaySearch, forecast that 3D TVs will rise to more than 50 percent of revenues and hit the 100-million unit sales mark by 2014.





BUT…

- **3D is hard on eyes and brain in long duration although exciting in short burst**

- **What if family members/friends watch the same movie, some prefer 3D and others 2D?**
  - 3D contents not watchable without glasses on existing 3D displays

- **The TPVM technology allows the same display to simultaneously exhibit 3D and 2D versions of a movie**
  - 3D with glasses
  - 2D without

2D and 3D Concurrent Viewing (II)

For example, in 3D TV of 240Hz refresh rate

$\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ --- $M = 4$ basis images

$\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2$ --- center, left and right view,

$$\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2 = \left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\right) \begin{pmatrix} 1/4 & w_{11} & w_{21} \\ 1/4 & w_{12} & w_{22} \\ 1/4 & w_{13} & w_{23} \\ 1/4 & w_{14} & w_{24} \end{pmatrix}$$

$$\mathbf{w}_1 = \left(w_{11}, w_{12}, w_{13}, w_{14}\right)^T$$

$$\mathbf{w}_2 = \left(w_{21}, w_{22}, w_{23}, w_{24}\right)^T$$ --- modulation vectors

A Practical 3D/2D Approach

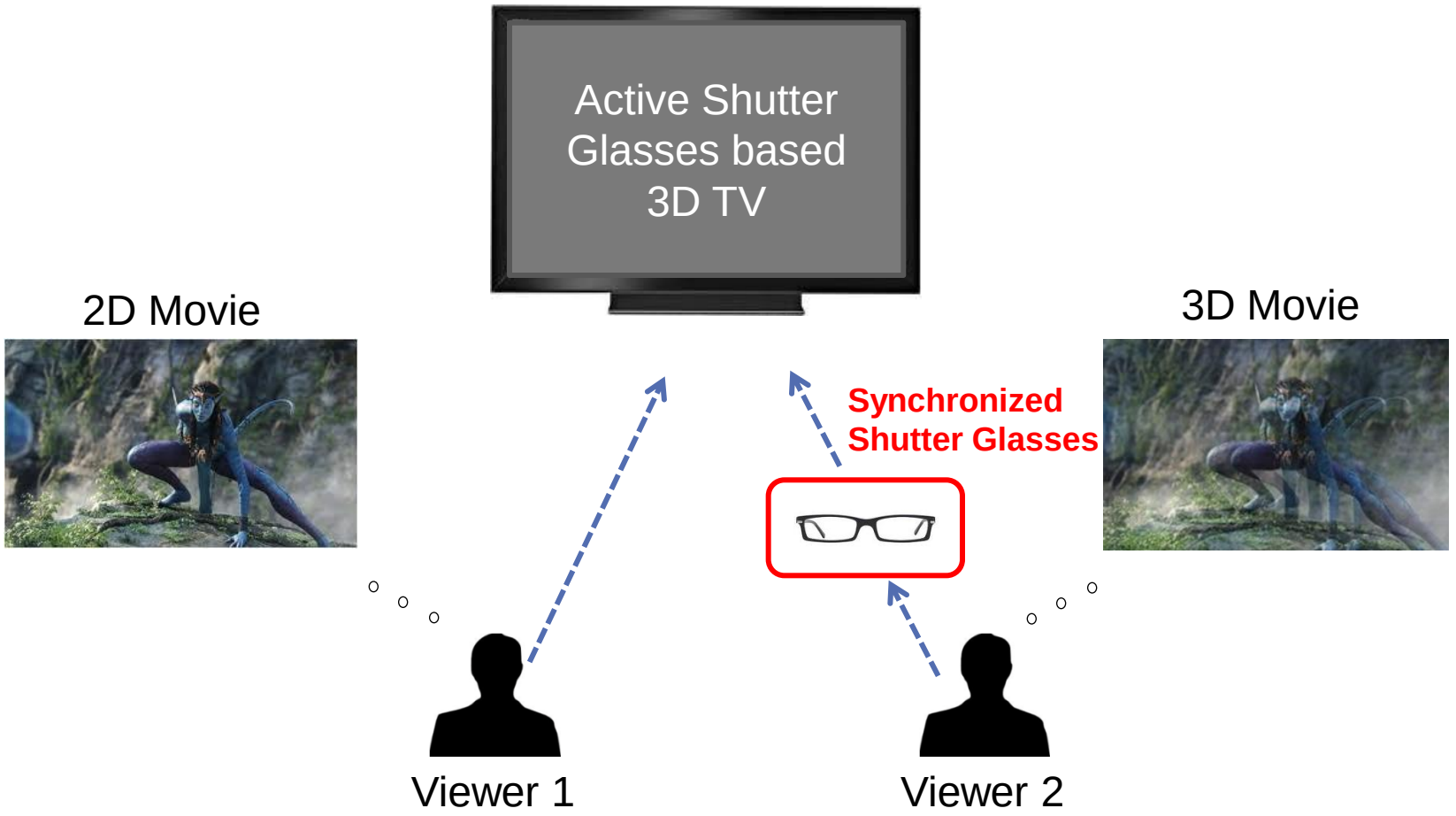
Solving the following linear least-square problem

$$\min_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4,\alpha} \left\{ \left\| \mathbf{x}_1 - \mathbf{y}_1 \right\| + \left\| \mathbf{x}_3 - \mathbf{y}_2 \right\| + \lambda \left\| \alpha \mathbf{y}_0 - \mathbf{x}_1 - \mathbf{x}_2 - \mathbf{x}_3 - \mathbf{x}_4 \right\| \right\}$$

$$\text{subject to } 0 \le \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \le 1$$

where the Lagrangian multiplier λ is used to balance the reconstruction errors in synthesized 2D and 3D views.

3D/2D Concurrency



Seeing is not Believing

TPVM can give multiple persons in the same physical presence very different visual percepts.

Purposefully make the normal view $\mathbf{y}_0$ visually and semantically unrelated to a shale view $\mathbf{y}_k$.

Applications of TPVM in information security and privacy, electronic payment, customized annotation, steganography and etc.

Perceptual bifurcation

A special case of NMF, assuming $M=2$

$$\left(\mathbf{y}_0, \mathbf{y}_1\right) = \left(\mathbf{x}_1, \mathbf{x}_2\right)\begin{pmatrix} 1 & w_1 \\ 1 & w_2 \end{pmatrix}$$

$\mathbf{y}_0$ the cover image;

$\mathbf{y}_1$ the secrete images;

$\mathbf{x}_1, \mathbf{x}_2$ basis images;

$\mathbf{w}_1, \mathbf{w}_2$ modulation weight vectors;

For the best legibility of the secrete message , we let $\mathbf{x}_1 = \mathbf{y}_1$, i.e.,

$$w_1 = 1, w_2 = 0$$

Make the normal view appear a random noise image $\mathbf{n}$ by setting

$$\mathbf{x}_2 = \mathbf{n} - \mathbf{x}_1$$

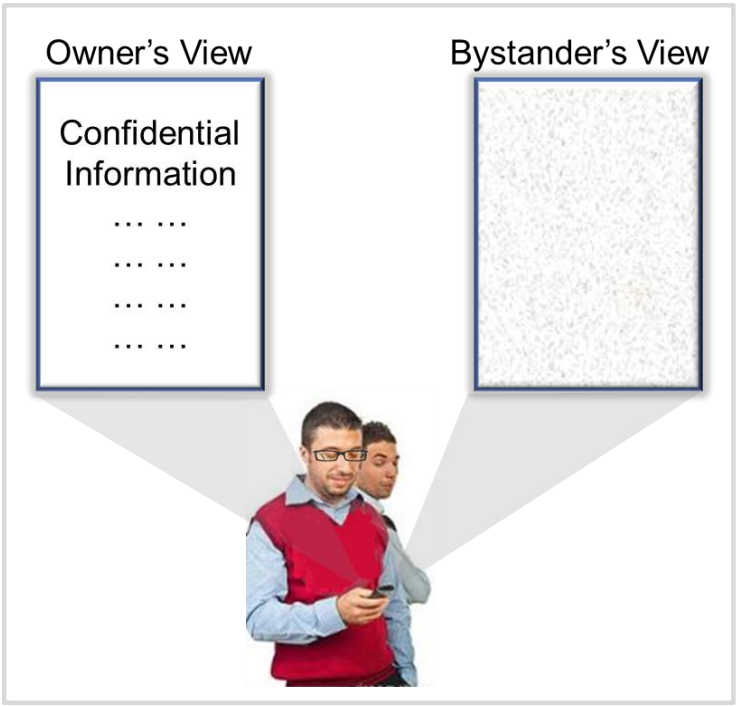
Lots of Applications!

Two classes of applications, categorized by the function of the normal view.

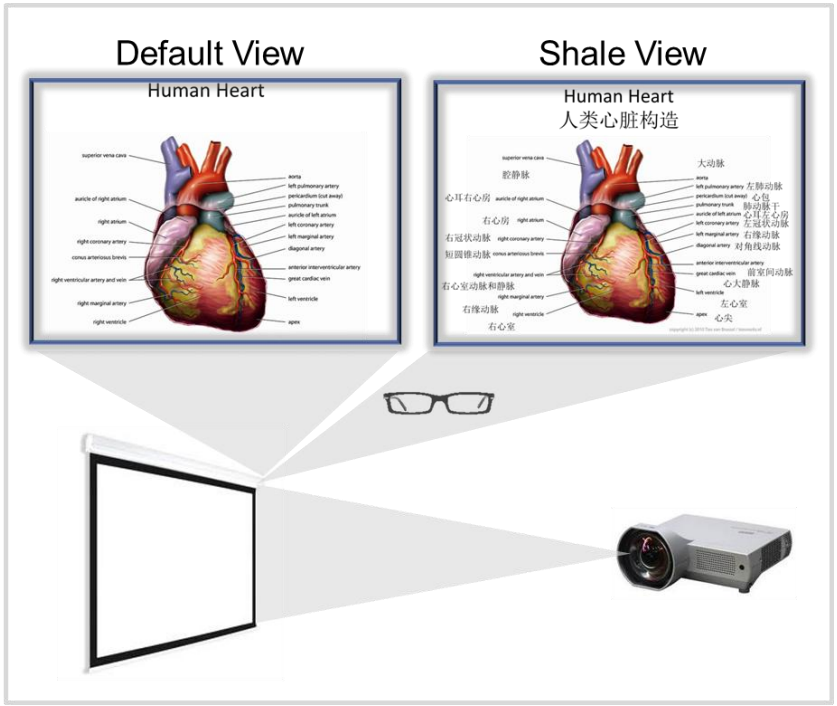
1. The normal view $\mathbf{y}_0$ plays a role similar to the covertext in steganography; The shale view $\mathbf{y}_k$ carries secrete messages that are only readable through an encrypted modulated viewing device.

2. The normal view is not to conceal but as a default view for majority or casual viewers, whereas the shale view is for some users of special needs who have to share the display with others.

Seeing is not Believing examples



Security



Annotation

Privacy Screens

Privacy protection in public areas

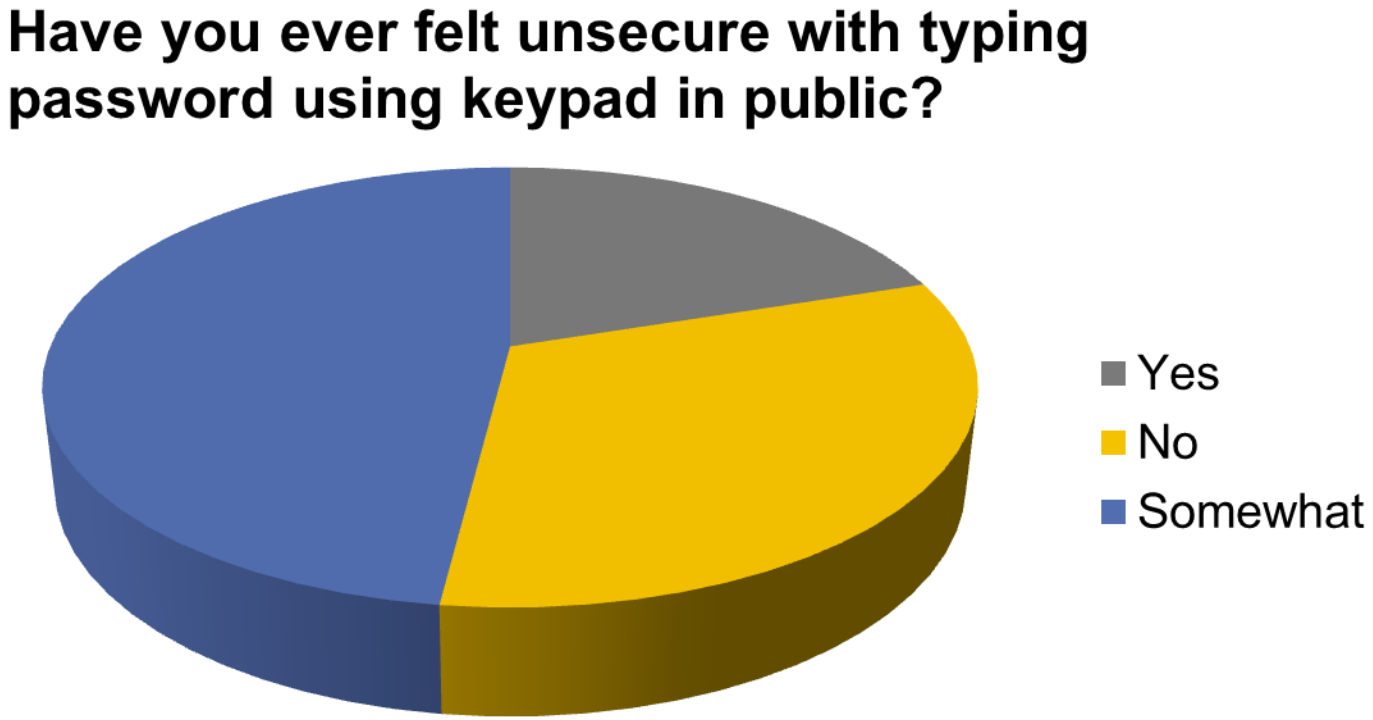
Laptop

Tablet

PDA

Smart Phone

Results of a 2010 Study



Over 70% of users think existing keypads are insecure

Invisible Keypads

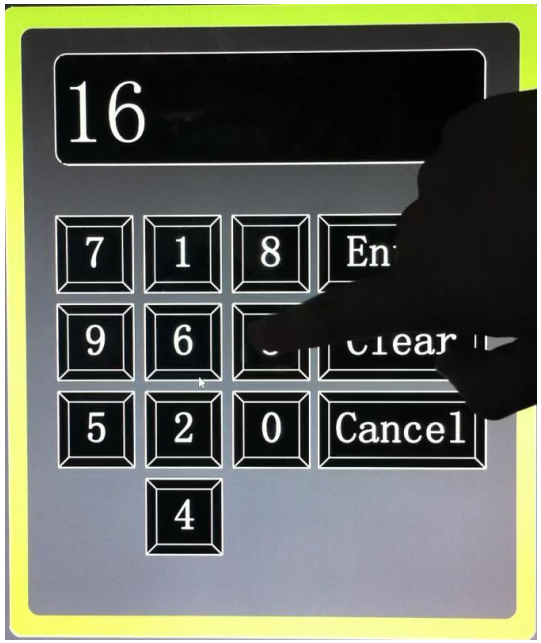
e-payment/transaction devices and solutions

POS

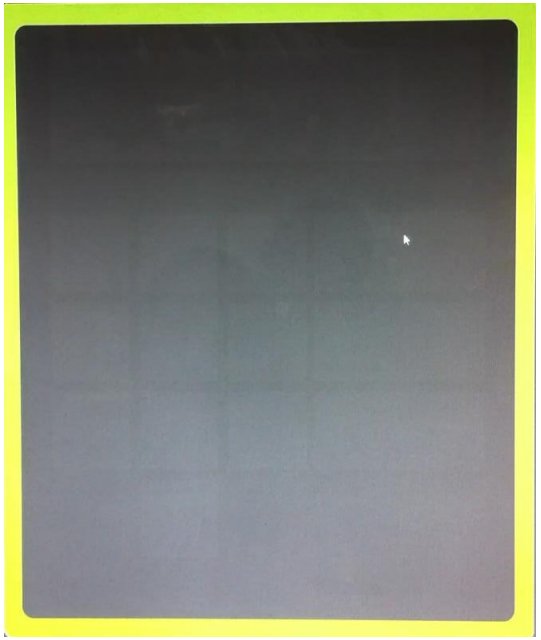
ATM

Snapshots of Quasion's prototype of invisible keypads

User's view

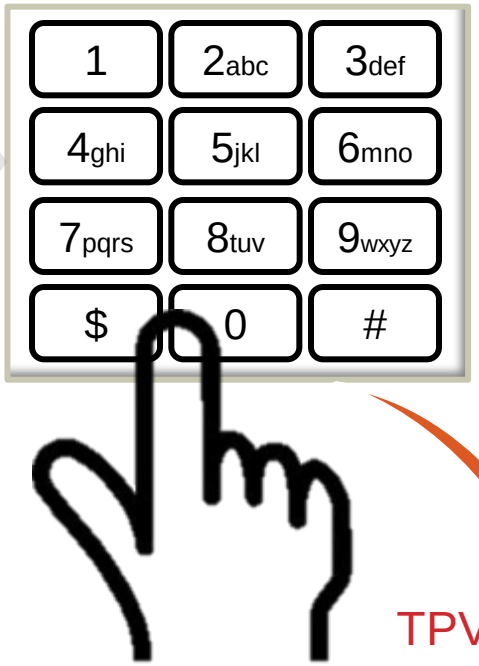


Bystander's view



INSECURE

Visible fixed keys

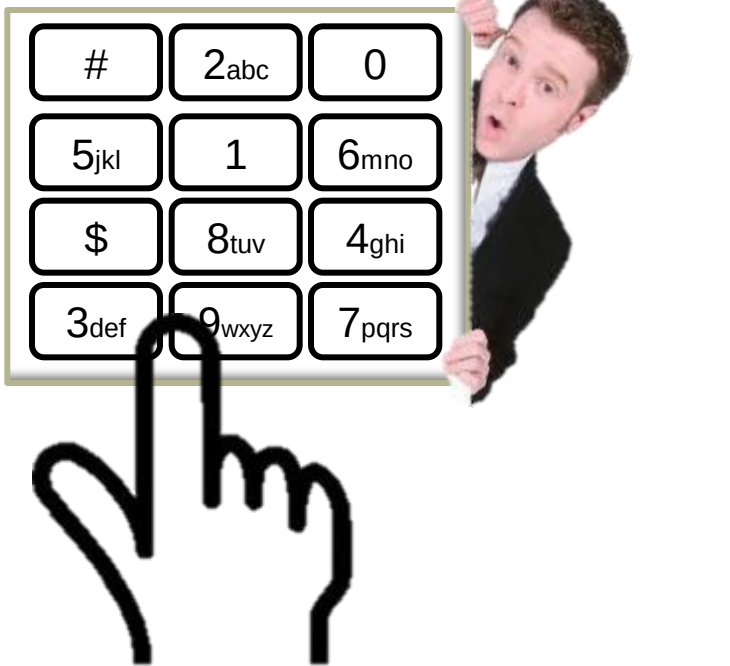


Advantages over existing technology

TPVM Technology

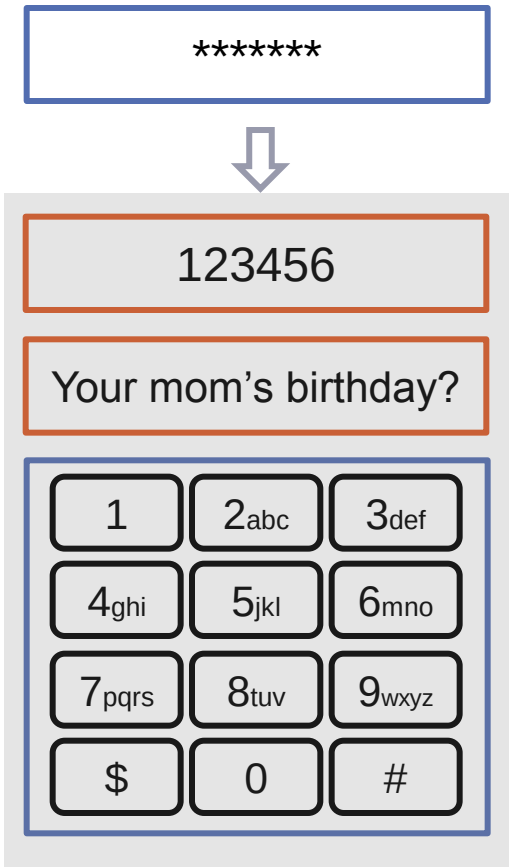
SECURE

Invisible randomized



More user friendly in addition to invisibility

Live password echo and reminder



Personalized keypad

Multilingual keypad

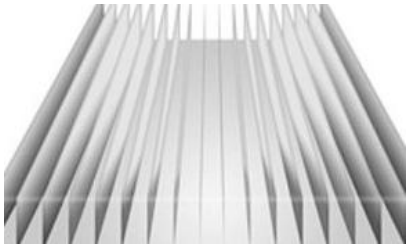
Use of colors and icons also possible.

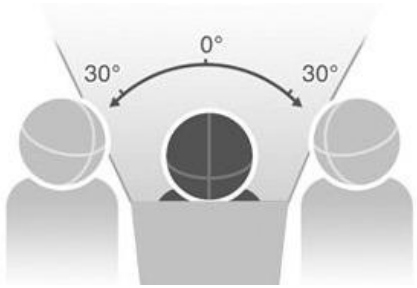
Competitor For Privacy Screens

Blocking/diverting lights for privacy protection

3M Privacy Filter using Advanced Light Control Film

- **Based on the microlouver technology.**
- **Fixed viewing angle (≈60°).**





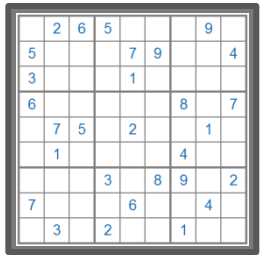
Drawbacks:

- **No guarantee of security (reading from behind).**
- **Geometrical distortion, glare and other artifacts.**

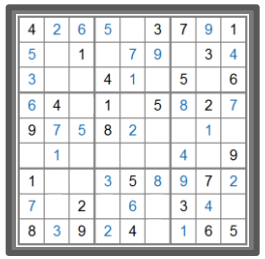
Immersive Gamming

Example : A Sudoku Game

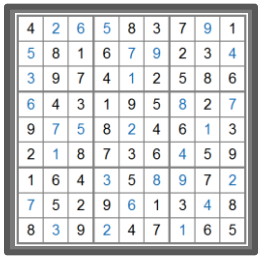
Problem



Problem
with Hint



Solution

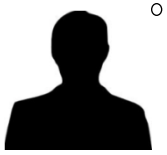


Player 1

Without Glasses
Natural experience

Player 2

Aided by Glasses
Hint of the game
Augmented experience

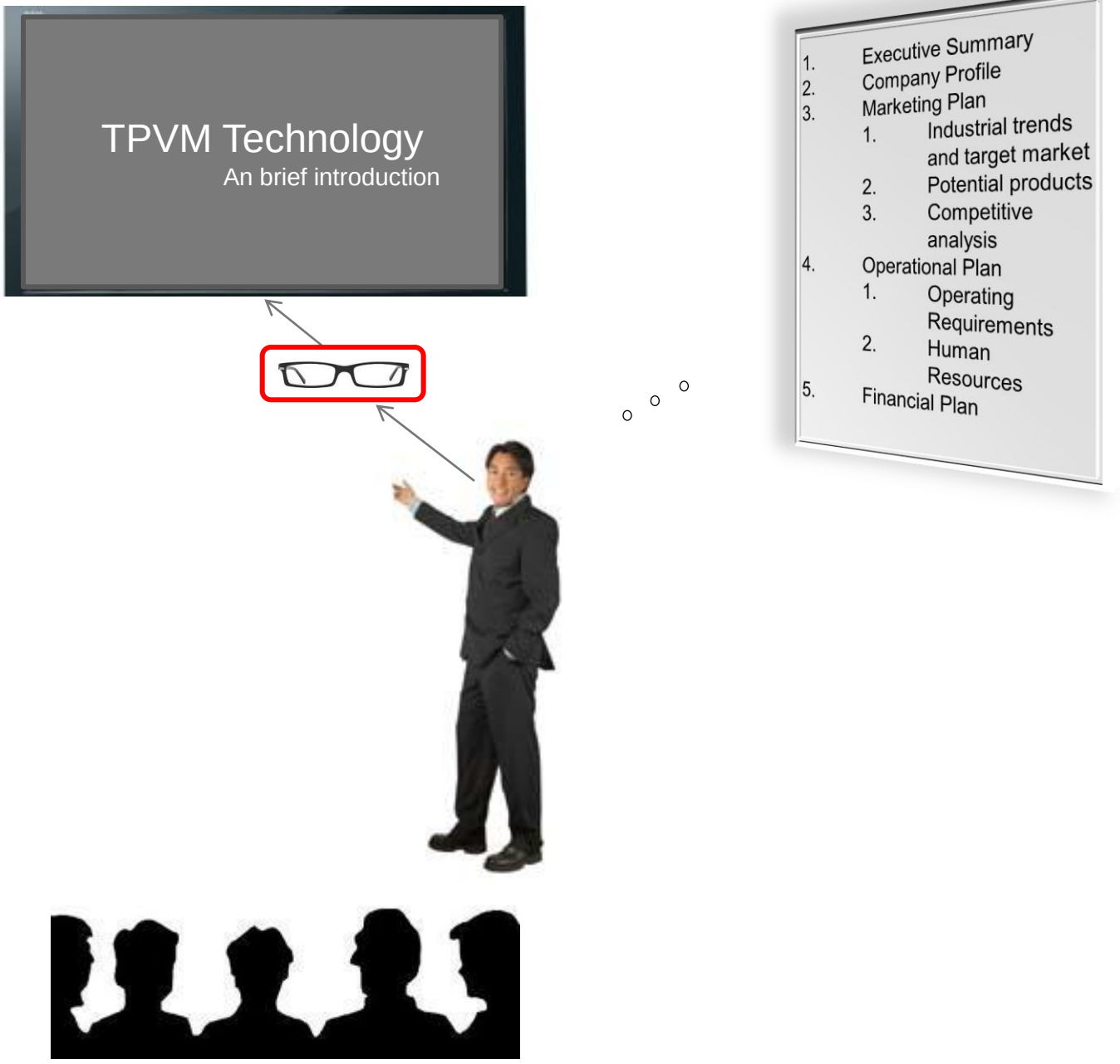


Observer

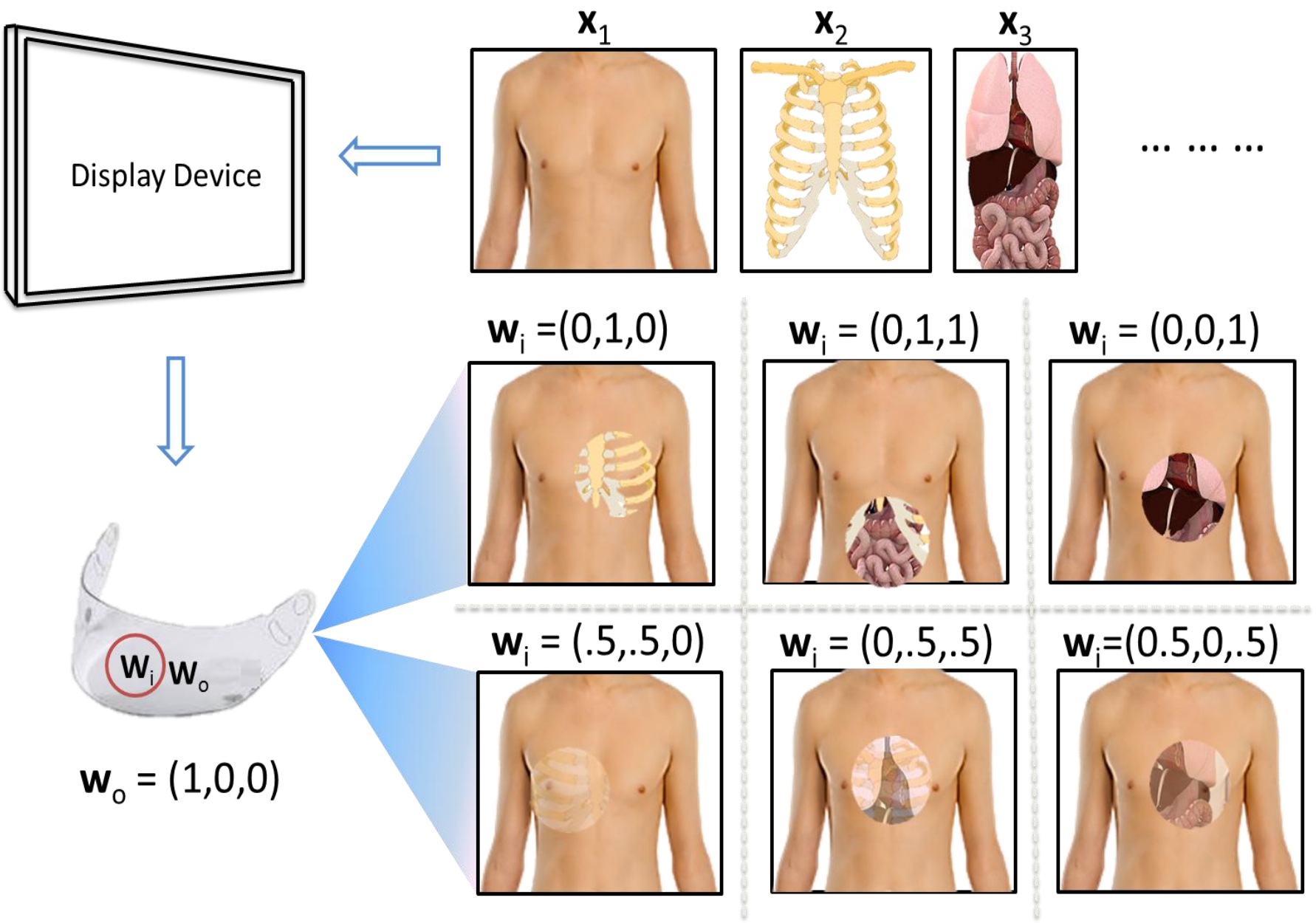
Aided by Glasses
God's view
Immersive experience

Shared Screen

Script reading for presenters (Powerpoint plugin)

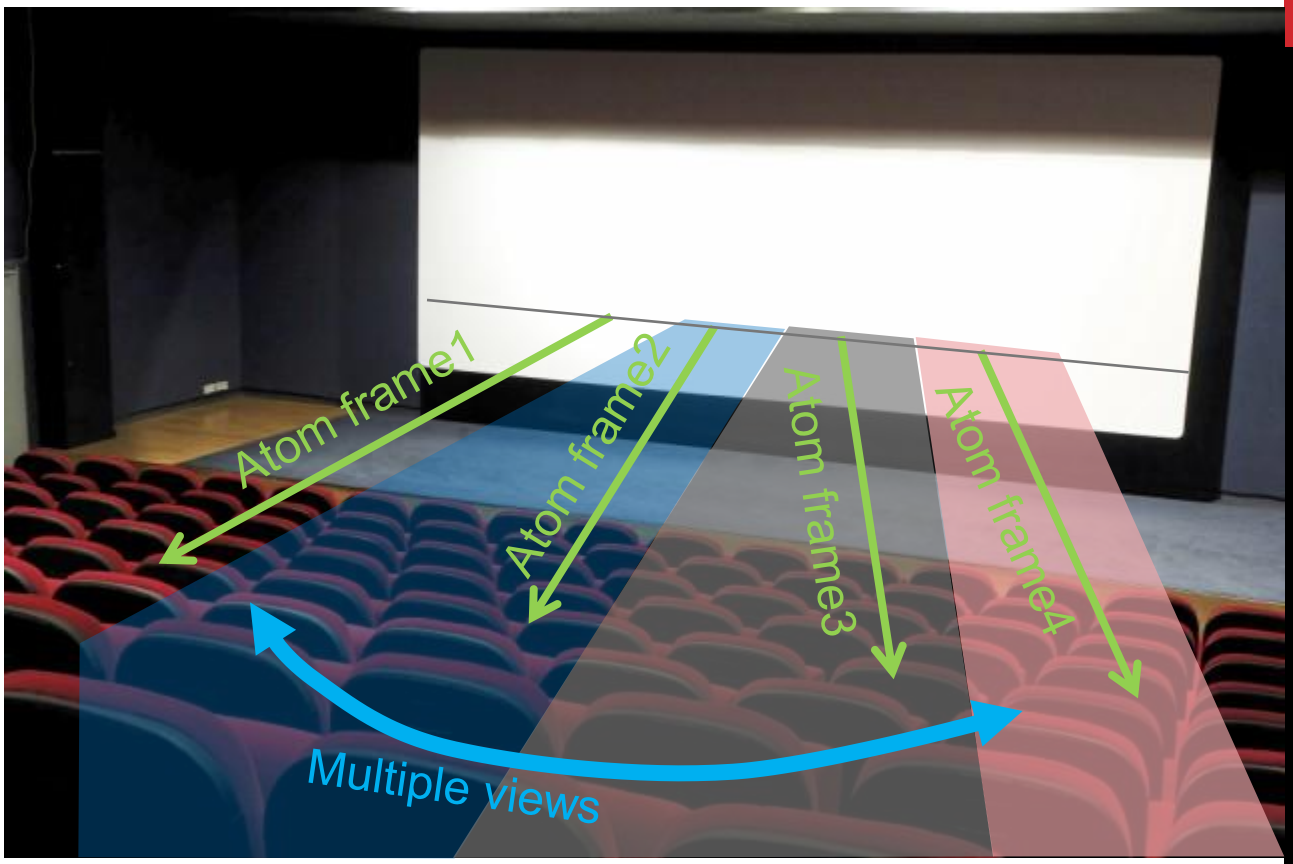


Visualization/Augmented Reality

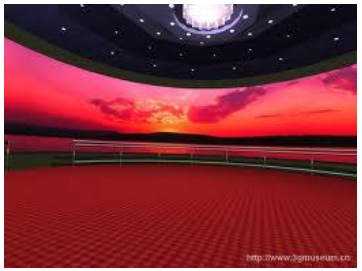


Interactive 3D cinema

- Multiple views through linear combination of atom frames
- Multiple view depths
- Virtual reality
- Immersive 3D viewing experience



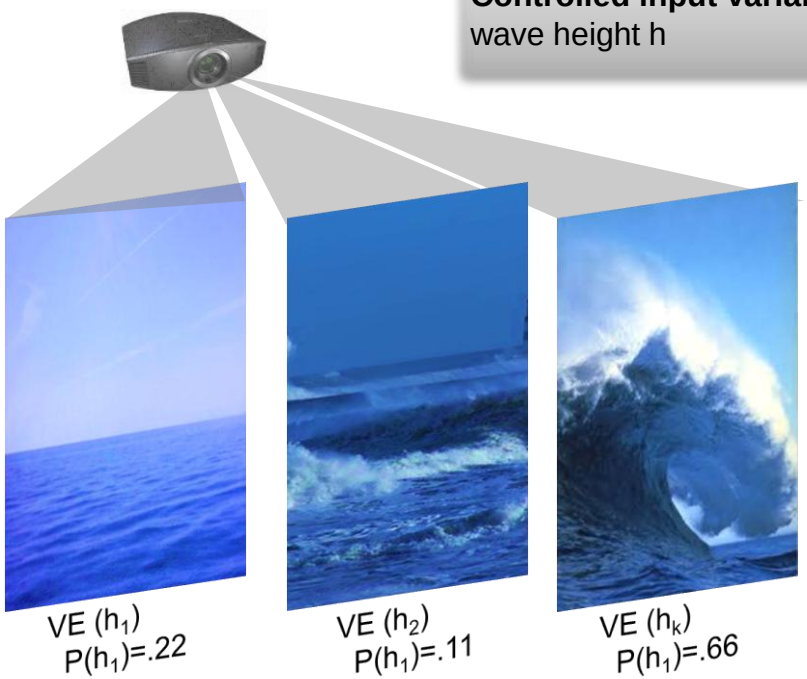
- Cinerama
- Panoramic movie
- Circular-screen movie



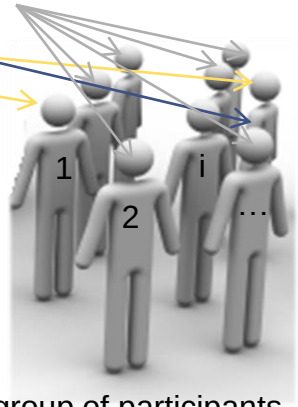
Psychophysical Experiments

TPVM image formation

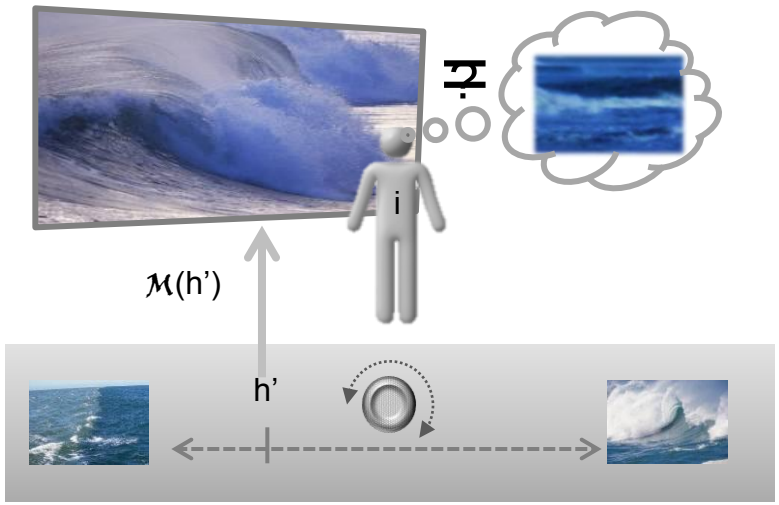
Controlled input variable:
wave height h



Physiological outputs:
heart rate: r_1, r_2,…, r_i
blood pressure: b_1, b_2,…, b_i
skin conductance: c_1, c_2,…, c_i
…



group of participants



Simpler, Lighter, Less Expensive VR

TPVM delegates bulk of imaging computations from HMD to HSV.

In TPVM-based VR, users wear glasses not helmet

V.S.

TPVM greatly reduces the computation power and video memory bandwidth required by VR, as a small number of basis frames (precomputed for a given scene) are used to synthesize a range of different perspective views.

How did we get the idea?

Destructive

to

Constructive

Science vs. Magic