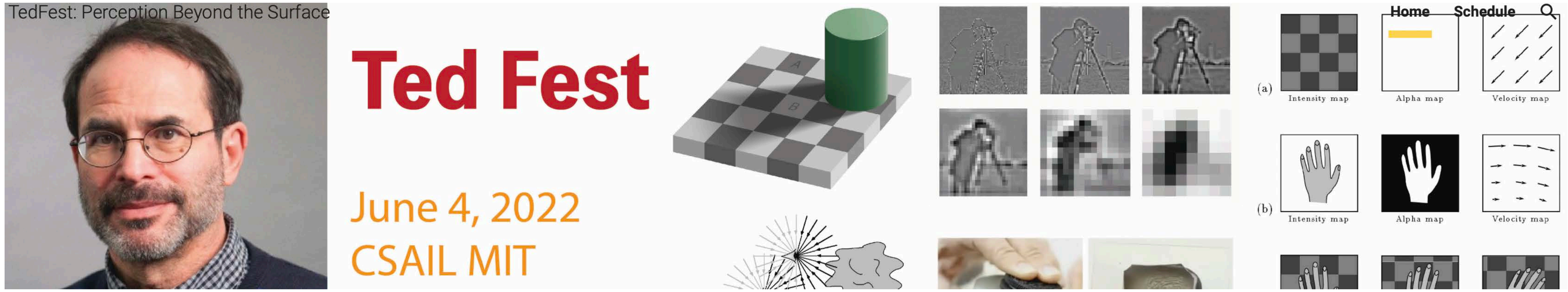
A self-absorbed look into how Ted Adelson and his ideas have enriched my scientific journey

Also how to generate informative line drawings from photographs

Frédo Durand, MIT



My main connection to Ted

I wasn't lucky enough to be his PhD student

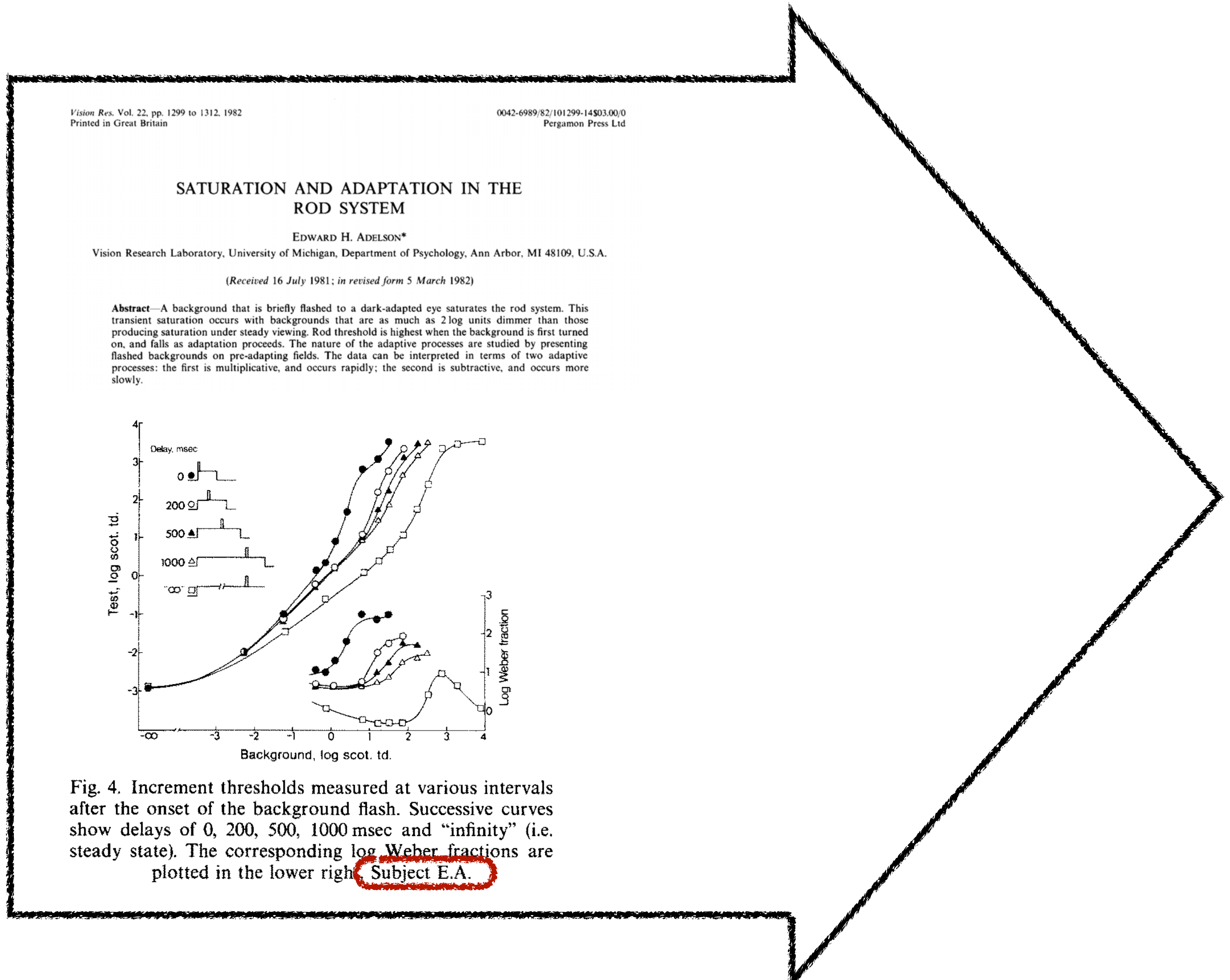
My main connection to Ted

But we have the same birthday (March 30)

After my PhD, I wanted to simulate visual adaptation

I quickly found that Ted’s work on rods was instrumental

- And learned that we must take into account perception of real scene as well as perception of an image



Vision Res. Vol. 22, pp. 1299 to 1312, 1982
Printed in Great Britain

0042-6989/82/101299-14$03.00/0
Pergamon Press Ltd

SATURATION AND ADAPTATION IN THE ROD SYSTEM

EDWARD H. ADELSON*
Vision Research Laboratory, University of Michigan, Department of Psychology, Ann Arbor, MI 48109, U.S.A.

(*Received* 16 *July* 1981; *in revised form* 5 *March* 1982)

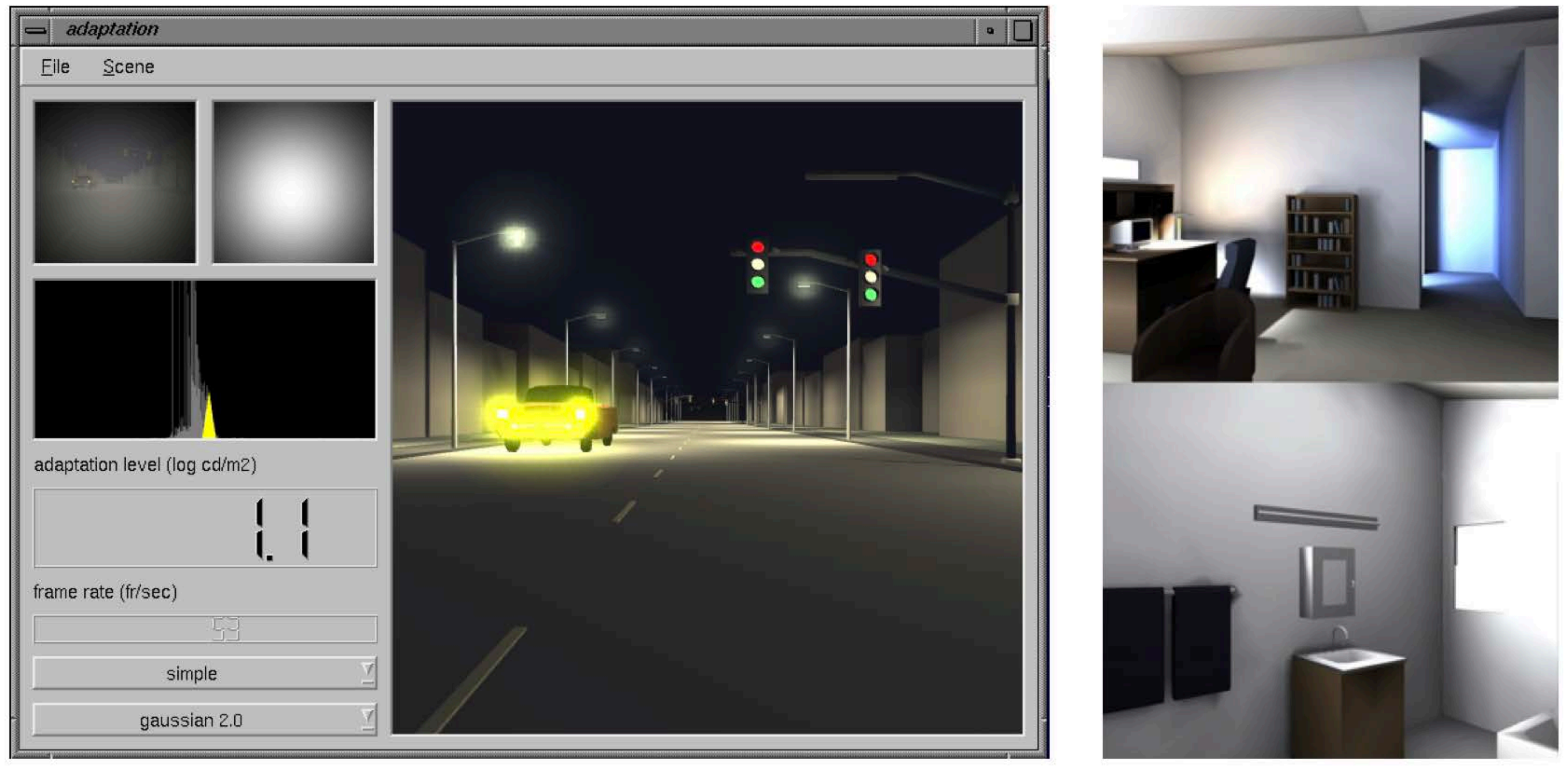
Abstract—A background that is briefly flashed to a dark-adapted eye saturates the rod system. This transient saturation occurs with backgrounds that are as much as 2 log units dimmer than those producing saturation under steady viewing. Rod threshold is highest when the background is first turned on, and falls as adaptation proceeds. The nature of the adaptive processes are studied by presenting flashed backgrounds on pre-adapting fields. The data can be interpreted in terms of two adaptive processes: the first is multiplicative, and occurs rapidly; the second is subtractive, and occurs more slowly.

Fig. 4. Increment thresholds measured at various intervals after the onset of the background flash. Successive curves show delays of 0, 200, 500, 1000 msec and "infinity" (i.e. steady state). The corresponding log Weber fractions are plotted in the lower righ Subject E.A.

Interactive Tone Mapping

Frédo Durand and Julie Dorsey

Laboratory for Computer Science
Massachusetts Institute of Technology
fredo@graphics.lcs.mit.edu, dorsey@lcs.mit.edu
http://www.graphics.lcs.mit.edu



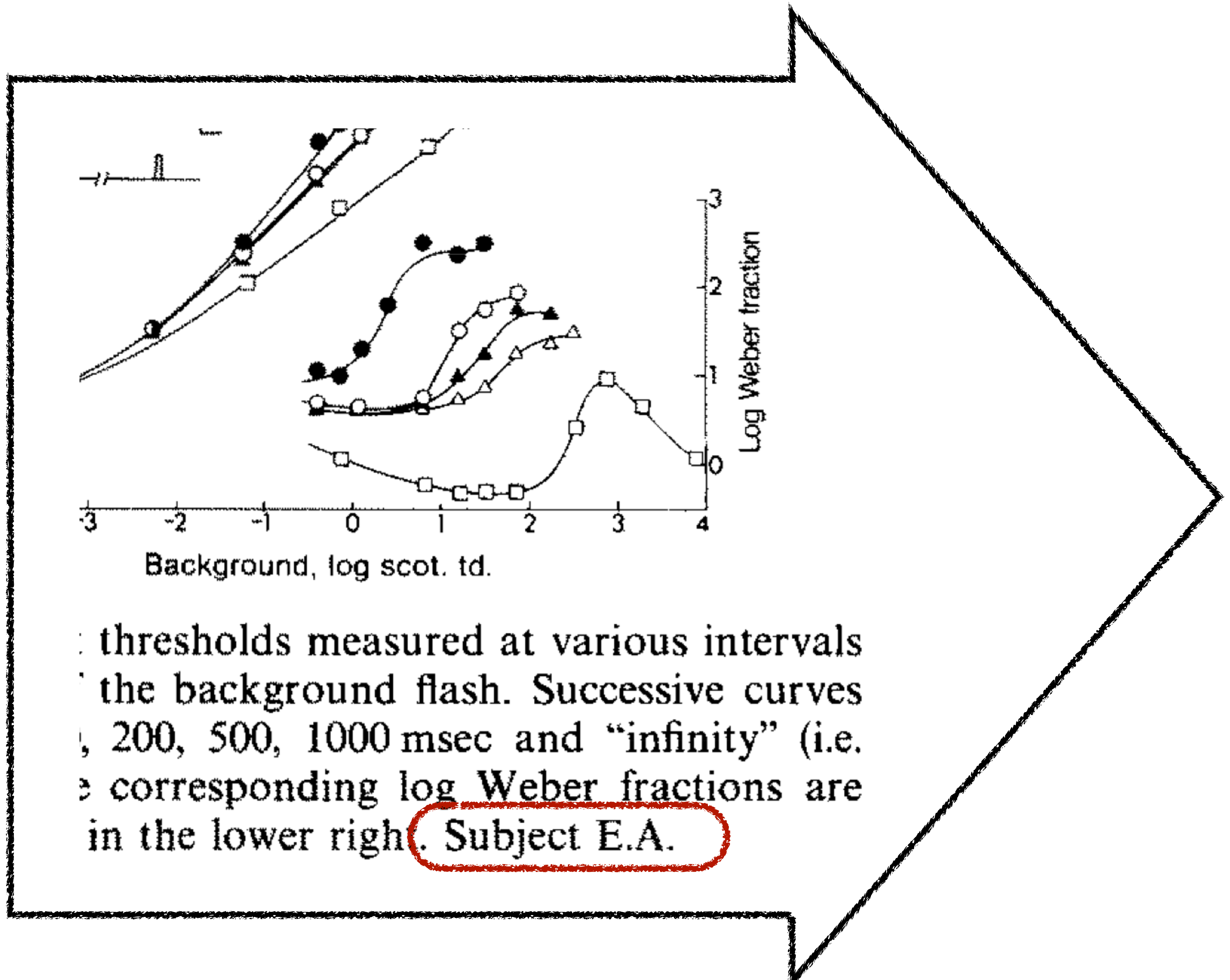
(a) (b)

Fig. 2. (a) Our interactive tone mapper for a street scene (70 ktri, 6Hz). The upper left window displays the scene with log colors. The window on the right is used to compute the normalization factor for the weighted average.Below is the histogram of scene luminance. (b) House scene (80 ktri, 6.6Hz). Top: Living room (1.4 log cd/m^2). Note the bluish lighting of the adjacent bathroom. Bottom: Bathroom after chromatic adaptation (1.8 log cd/m^2).

After my PhD, I wanted to simulate visual adaptation

I quickly found that Ted’s work on rods was instrumental

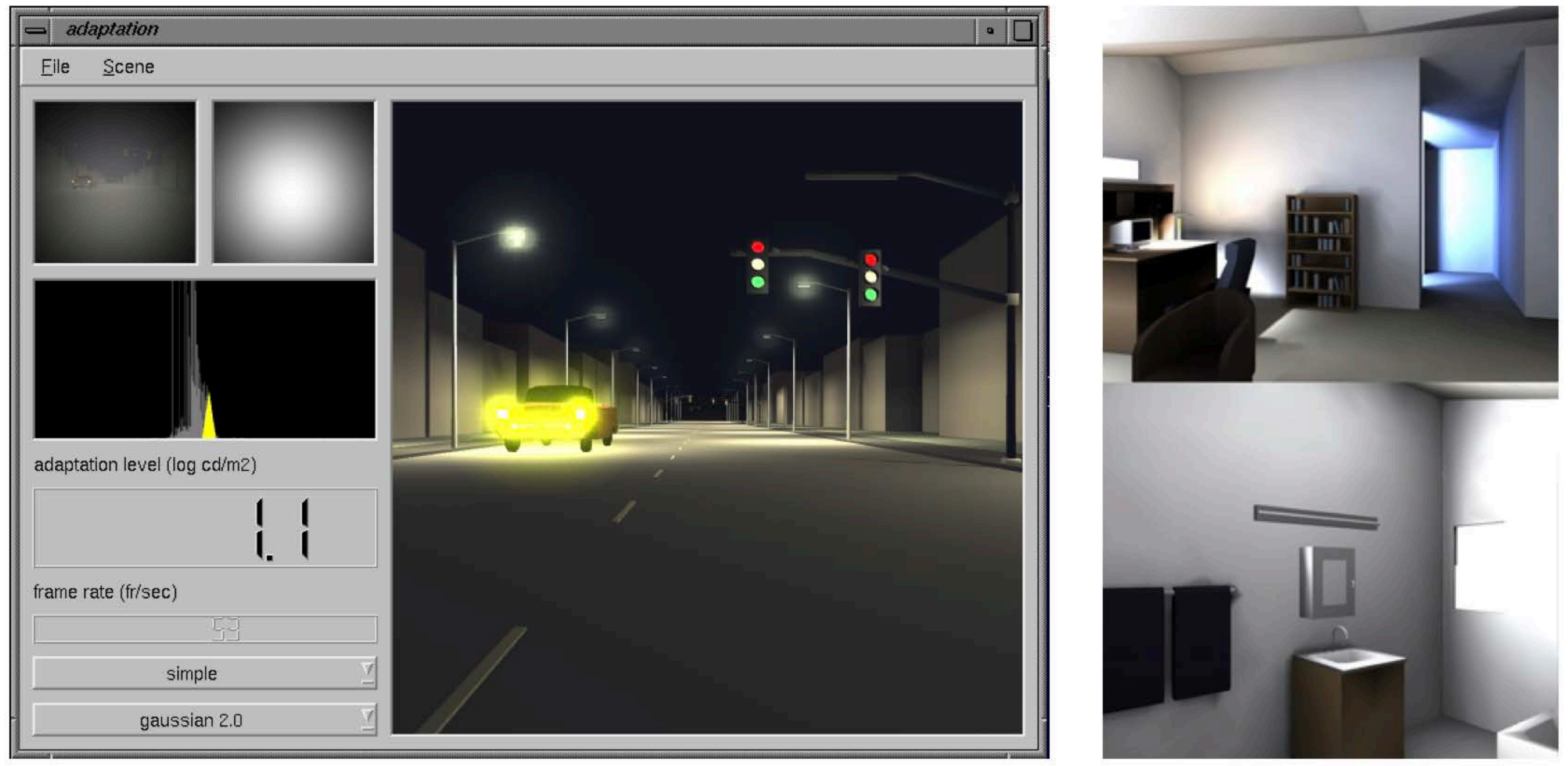
- And learned that we must take into account perception of real scene as well as perception of an image



thresholds measured at various intervals
the background flash. Successive curves
200, 500, 1000 msec and “infinity” (i.e.
corresponding log Weber fractions are
in the lower right. Subject E.A.

Interactive Tone Mapping

Frédo Durand and Julie Dorsey
Laboratory for Computer Science
Massachusetts Institute of Technology
fredo@graphics.lcs.mit.edu, dorsey@lcs.mit.edu
http://www.graphics.lcs.mit.edu



(a) (b)

Fig. 2. (a) Our interactive tone mapper for a street scene (70 ktri, 6Hz). The upper left window displays the scene with log colors. The window on the right is used to compute the normalization factor for the weighted average.Below is the histogram of scene luminance. (b) House scene (80 ktri, 6.6Hz). Top: Living room (1.4 log cd/m^2). Note the bluish lighting of the adjacent bathroom. Bottom: Bathroom after chromatic adaptation (1.8 log cd/m^2).

I attended Ted’s material appearance seminar

A transformative experience that inspired more than my work on materials

- The wonders of exploring what questions may hide
- In research, finding a good question is more than half of the work
- Interdisciplinary work is important and exciting
- There is a lot we can learn from art

Statistical Acquisition of Texture Appearance

Addy Ngan   Frédo Durand†

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology

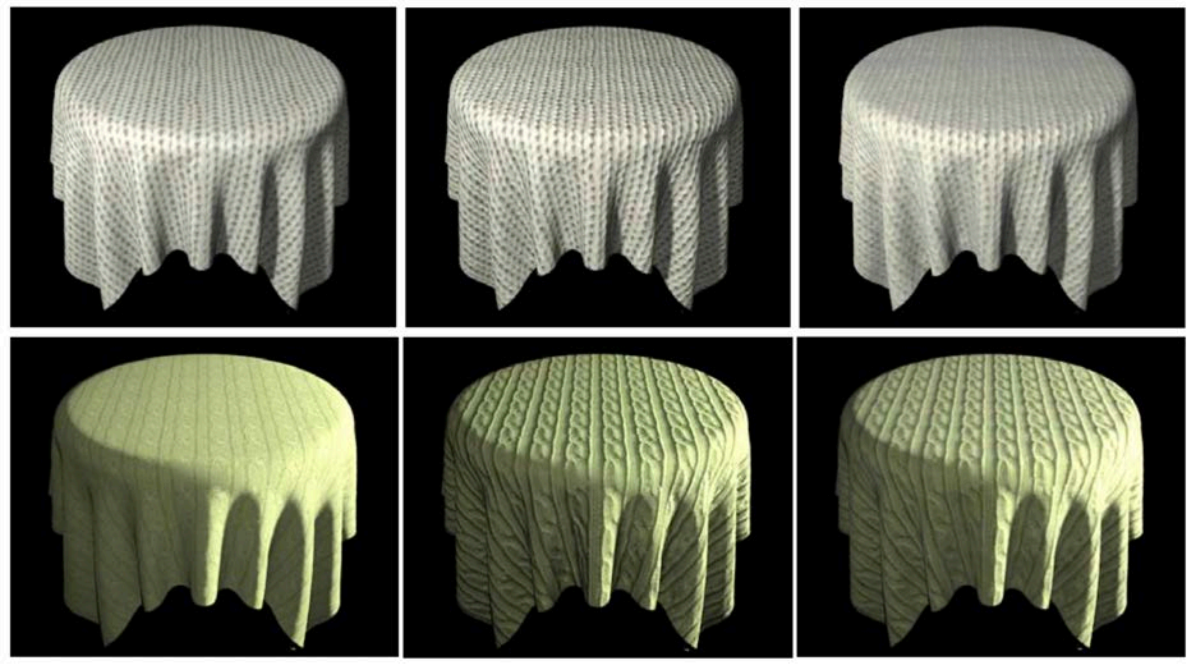


Figure 11: *Comparing approximations to the measured materials knitwear-1 and green-knitwear. First column: single texture modulated by acquired BRDF, second column: light-varying textures from top view, and third column: our reconstruction.*

Texture Transfer Using Geometry Correlation

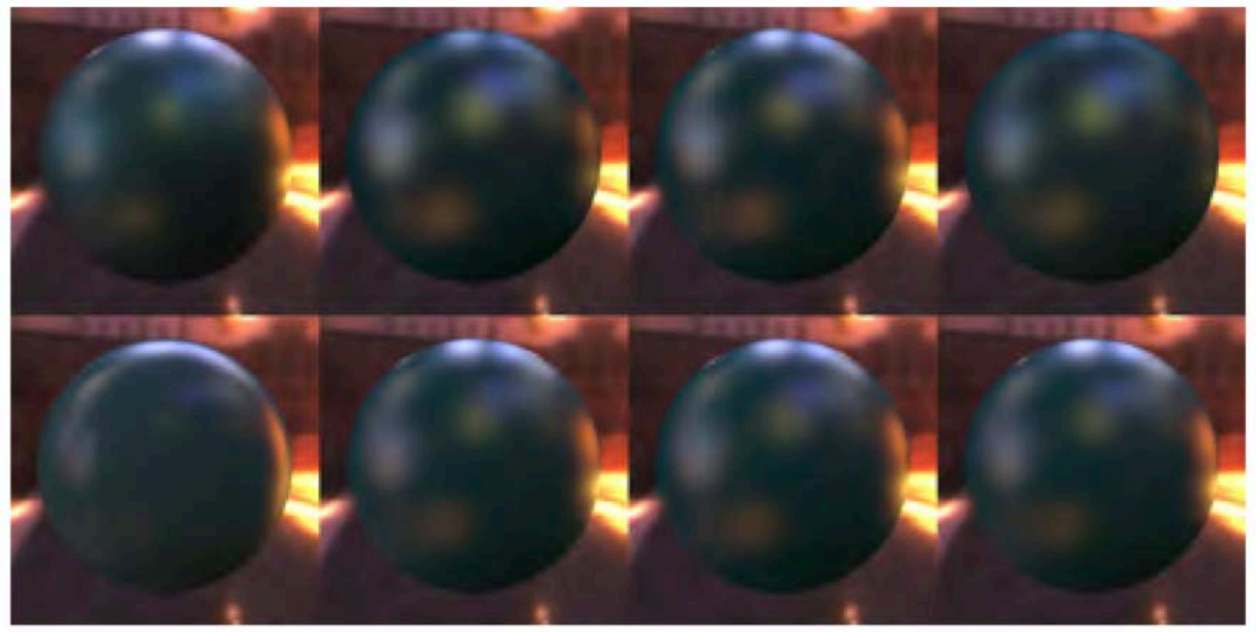
Tom Mertens, CSAIL – MIT† | Jan Kautz, University College London | Jiawen Chen, CSAIL – MIT | Philippe Bekaert, Hasselt University, EDM‡ – tUL§ | Frédo Durand, CSAIL – MIT

Figure 1: *The appearance of two texture-mapped models is transferred to a target model (the Bunny). We analyze the geometric features of the source and their correlation with texture. The source texture is transferred to the target mesh based on the correlation.*

Experimental Analysis of BRDF Models

Addy Ngan, Frédo Durand,† and Wojciech Matusik‡

MIT CSAIL   MERL

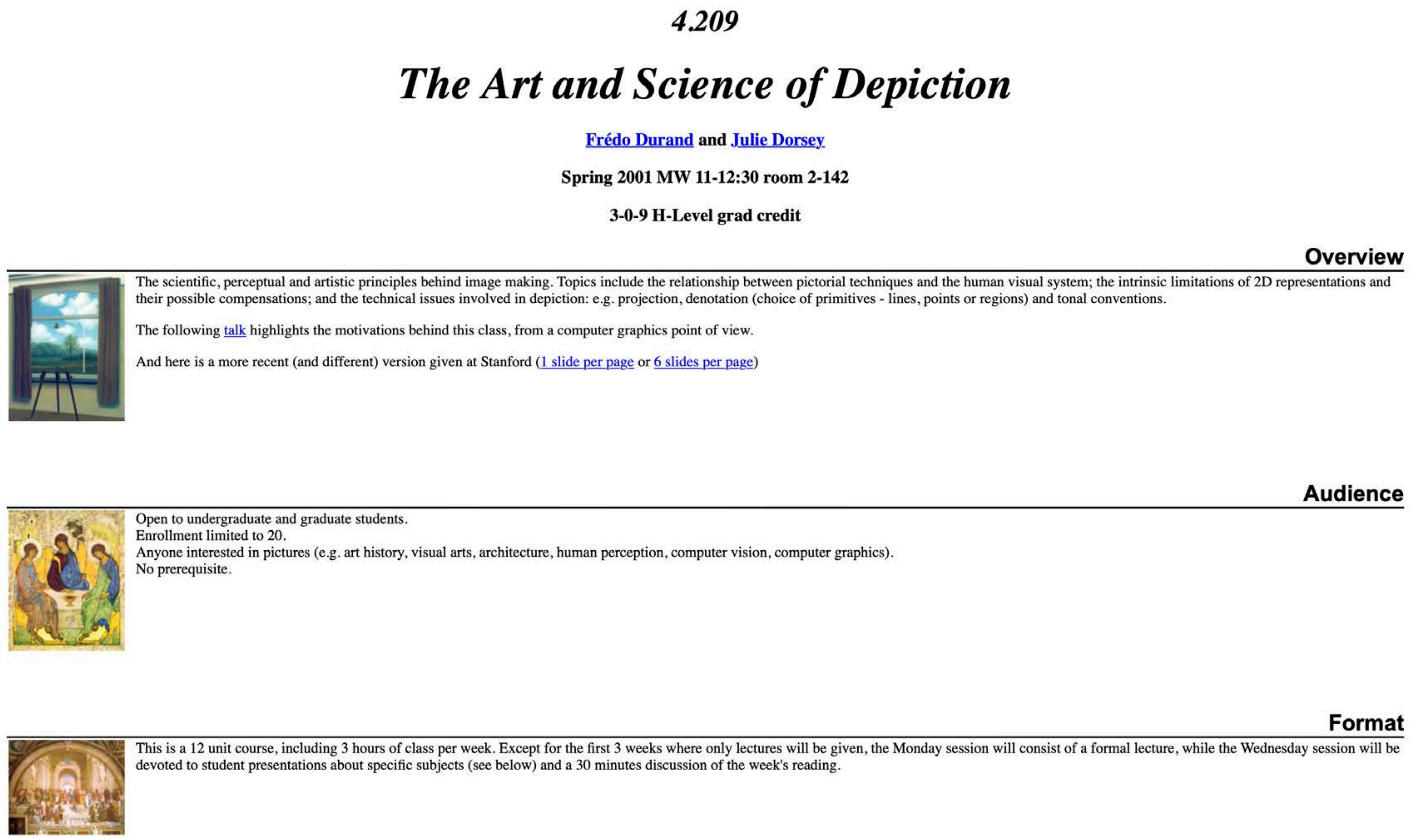


I was emboldened to explore how properties in images relate to our perception of the world

Class The Art and Science of Depiction

- https://people.csail.mit.edu/fredo/ArtAndScienceOfDepiction/

- I read a lot of design and art+science books. The majority of them showed the checkerboard illusion



4.209

The Art and Science of Depiction

Frédo Durand and Julie Dorsey

Spring 2001 MW 11-12:30 room 2-142

3-0-9 H-Level grad credit

Overview

The scientific, perceptual and artistic principles behind image making. Topics include the relationship between pictorial techniques and the human visual system; the intrinsic limitations of 2D representations and their possible compensations; and the technical issues involved in depiction: e.g. projection, denotation (choice of primitives - lines, points or regions) and tonal conventions.

The following talk highlights the motivations behind this class, from a computer graphics point of view.

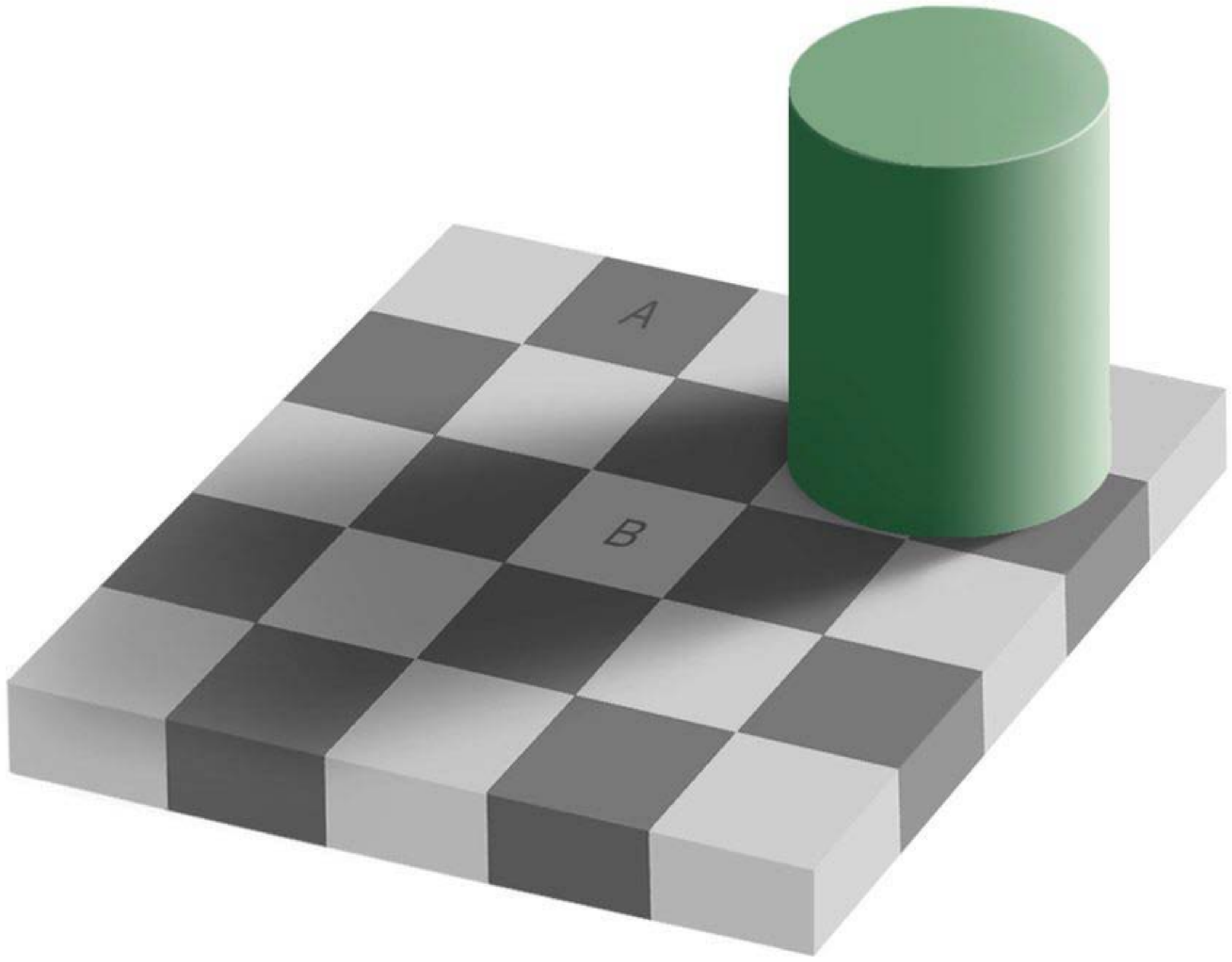
And here is a more recent (and different) version given at Stanford (1 slide per page or 6 slides per page)

Audience

Open to undergraduate and graduate students.
Enrollment limited to 20.
Anyone interested in pictures (e.g. art history, visual arts, architecture, human perception, computer vision, computer graphics).
No prerequisite.

Format

This is a 12 unit course, including 3 hours of class per week. Except for the first 3 weeks where only lectures will be given, the Monday session will consist of a formal lecture, while the Wednesday session will be devoted to student presentations about specific subjects (see below) and a 30 minutes discussion of the week's reading.



I argued that making images is an inverse of inverse problem and an optimization

Inspired in large part by Hermann von Helmholz

- And of course I had to include (questionably plagiarized) checkerboard illusions

An Invitation to Discuss Computer Depiction

Frédo Durand

Laboratory for Computer Science, MIT*

Abstract

This paper draws from art history and perception to place computer depiction in the broader context of picture production. It highlights the often underestimated complexity of the interactions between features in the picture and features of the represented scene. Depiction is not always a unidirectional projection from a 3D scene to a 2D picture, but involves much feedback and influence from the picture space to the object space. Depiction can be seen as a pre-existing 3D reality projected onto 2D, but also as a 2D pictorial representation that is superficially compatible with an hypothetic 3D scene. We show that depiction is essentially an optimization problem, producing the best picture given goals and constraints.

There is a variety of picture production purposes, resulting in very different contexts and specificities. We show the complexity and richness of depiction, and the discussion is independent of any implementation. Our main goal is to introduce a vocabulary that will make a principled discussion possible, and to raise questions rather than providing answers. We review and build upon visual arts and perception literature. We outline important issues of depiction that we use to discuss the field of non-photorealistic rendering, and more generally, computer depiction.

Computer graphics has long been defined as a quest to achieve *photorealism*. As it gets closer to this grail, the field realizes that there is more to images than realism alone. Non-photorealistic pictures can be more effective

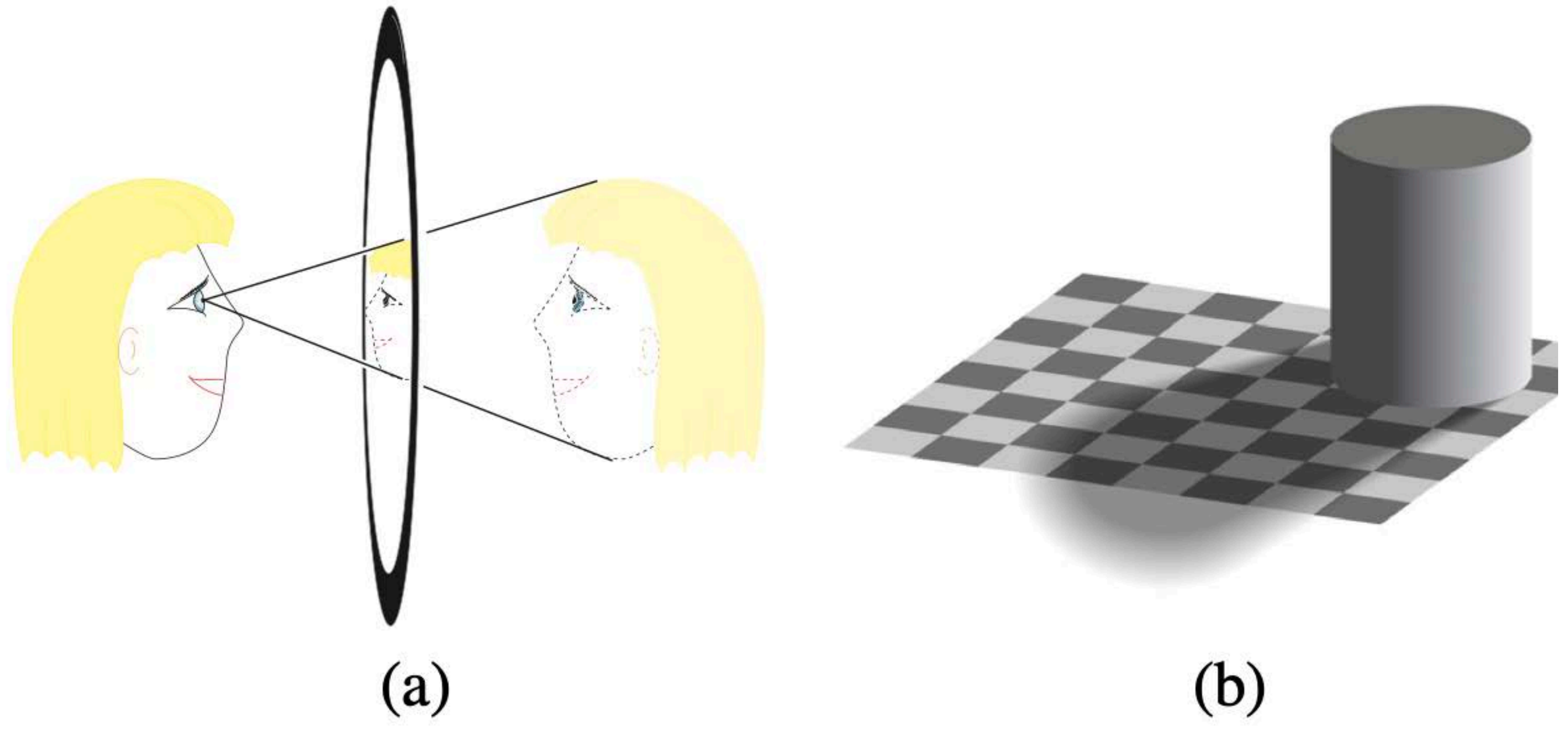


Figure 1: (a) Mirror illusion. The size of our reflection on the sur face of a mirror is half our size. (b) In this picture, the white cell in the shadow of the cylinder have the same grey level as the blac cells in full light. After an illusion by Ted Adelson.

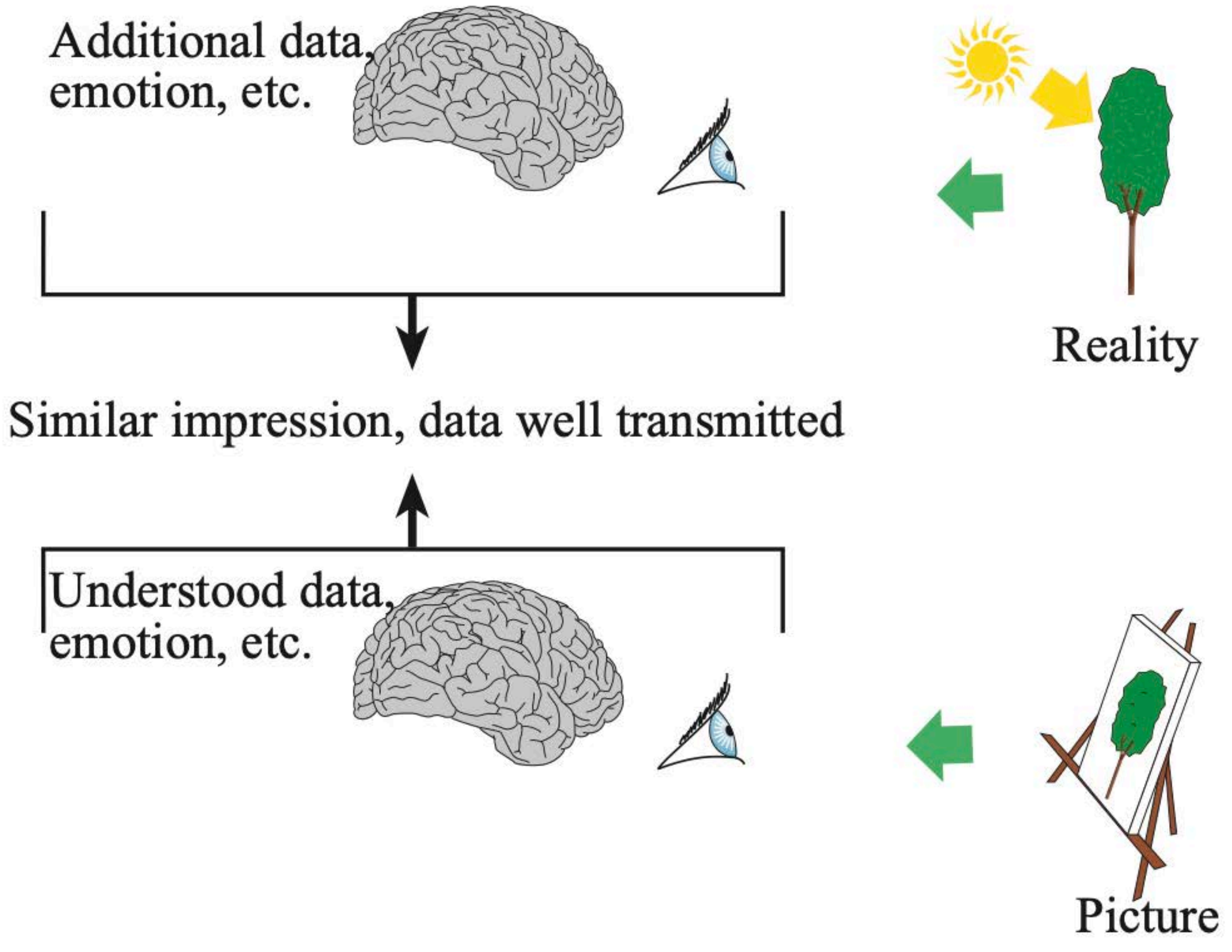


Figure 6: Depiction as the inverse of an inverse problem.

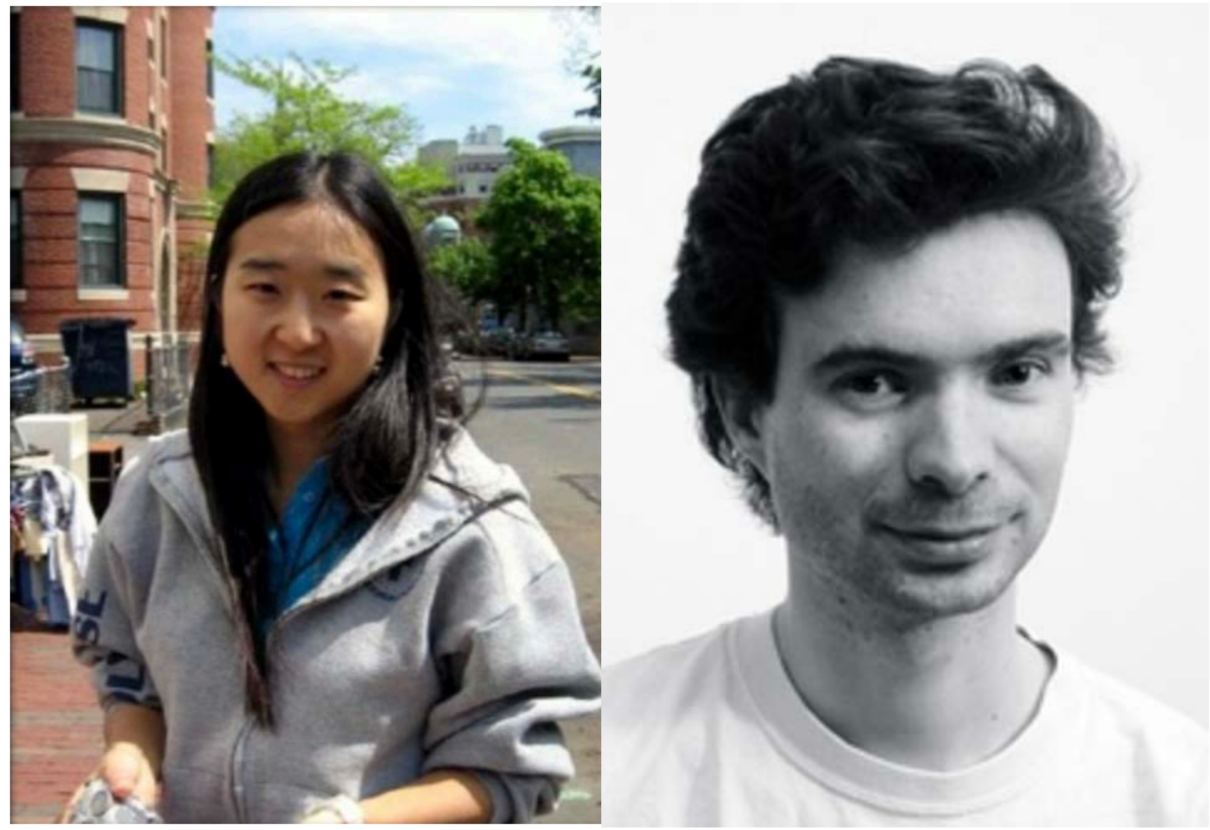
Then I had to choose a job in 2002

Then I had to choose a job in 2002

Then I had to choose a job in 2002

Photographic style transfer

Based on local texture energy



Early vision and texture perception

James R. Bergen* & Edward H. Adelson†

* SRI David Sarnoff Research Center, Princeton, New Jersey 08540, USA
† Media Lab and Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

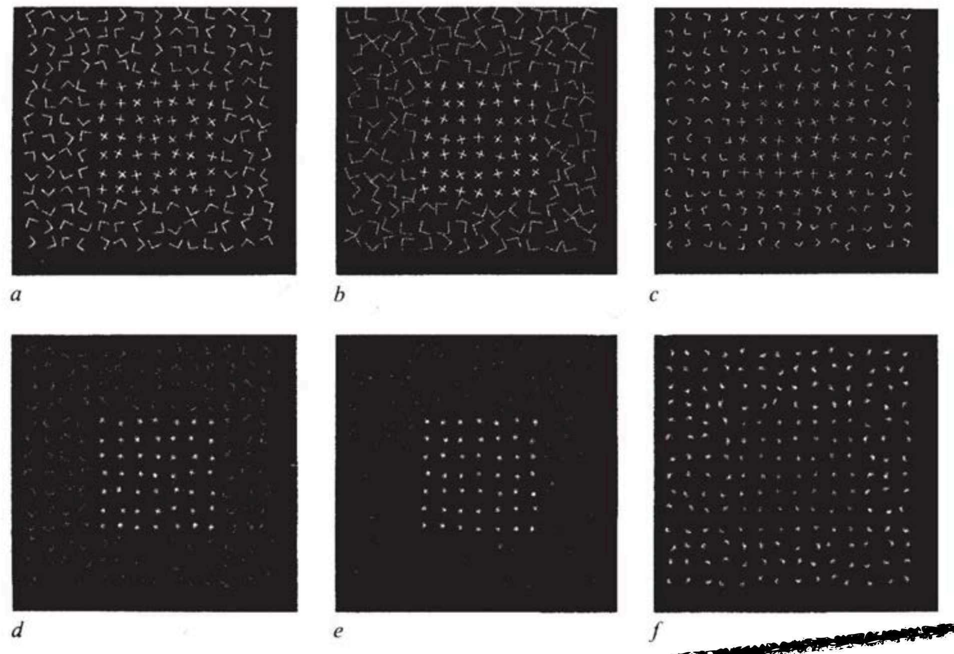
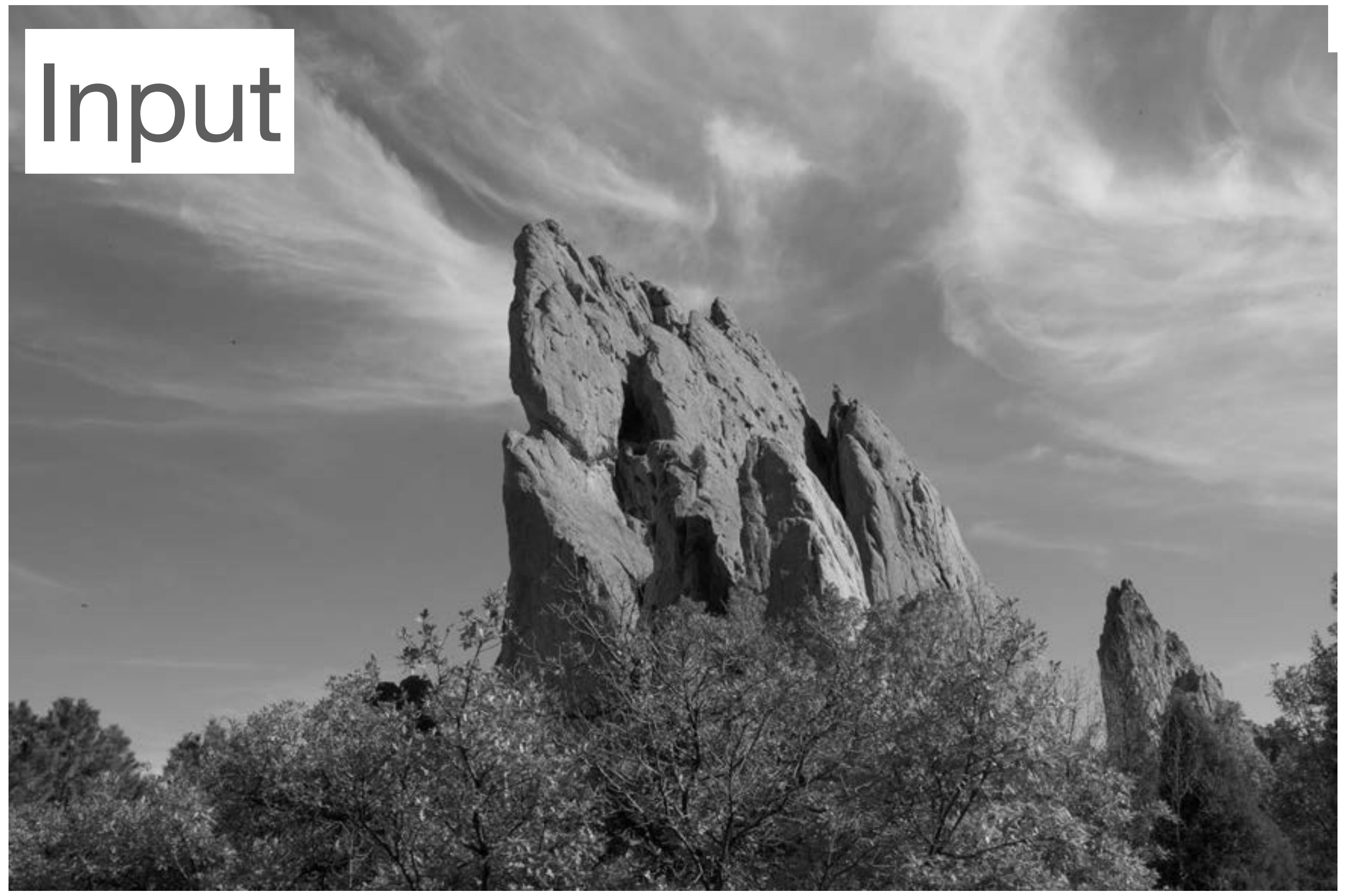
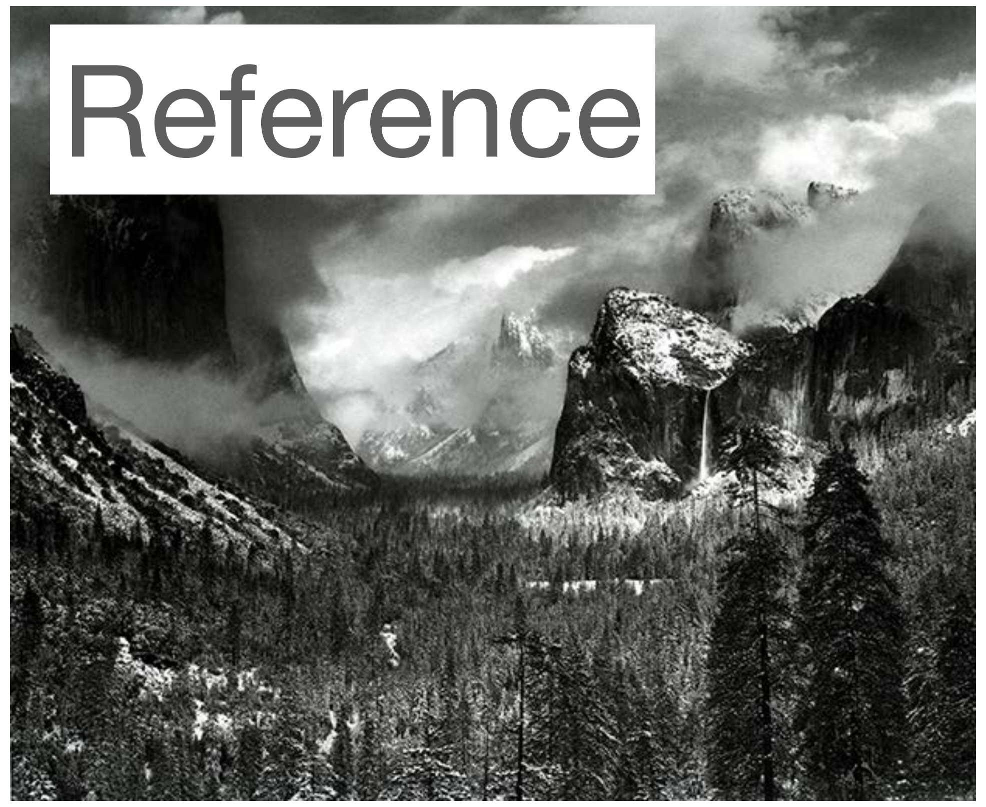


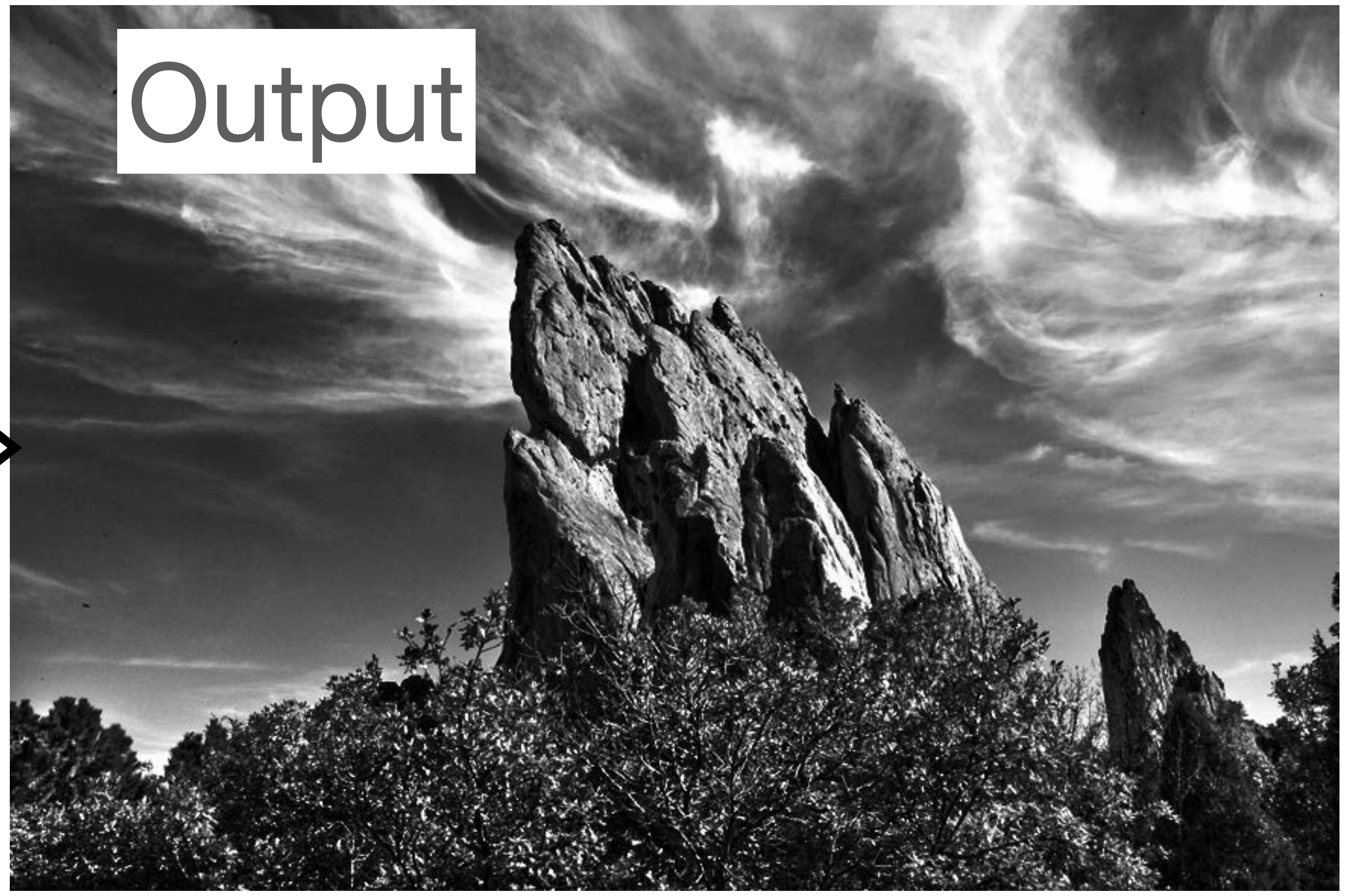
Fig. 1 *Top row*, Textures consisting of Xs within a texture composed of Ls. The micropatterns are placed at random orientations on a randomly perturbed lattice. *a*, The bars of the Xs have the same length as the bars of the Ls. *b*, The bars of the Ls have been lengthened by 25%, and the intensity adjusted for the same mean luminance. Discriminability is enhanced. *c*, The bars of the Ls have been shortened by 25%, and the intensity adjusted for the same mean luminance. Discriminability is impaired. *Bottom row*, The responses of a size-tuned mechanism. *d*, response to image *a*; *e*, response to image *b*; *f*, response to image *c*.

Two-scale Tone Management for Photographic Look

Soonmin Bae Sylvain Paris Frédo Durand
Computer Science and Artificial Intelligence Laboratory
Massuchusetts Institute of Technology

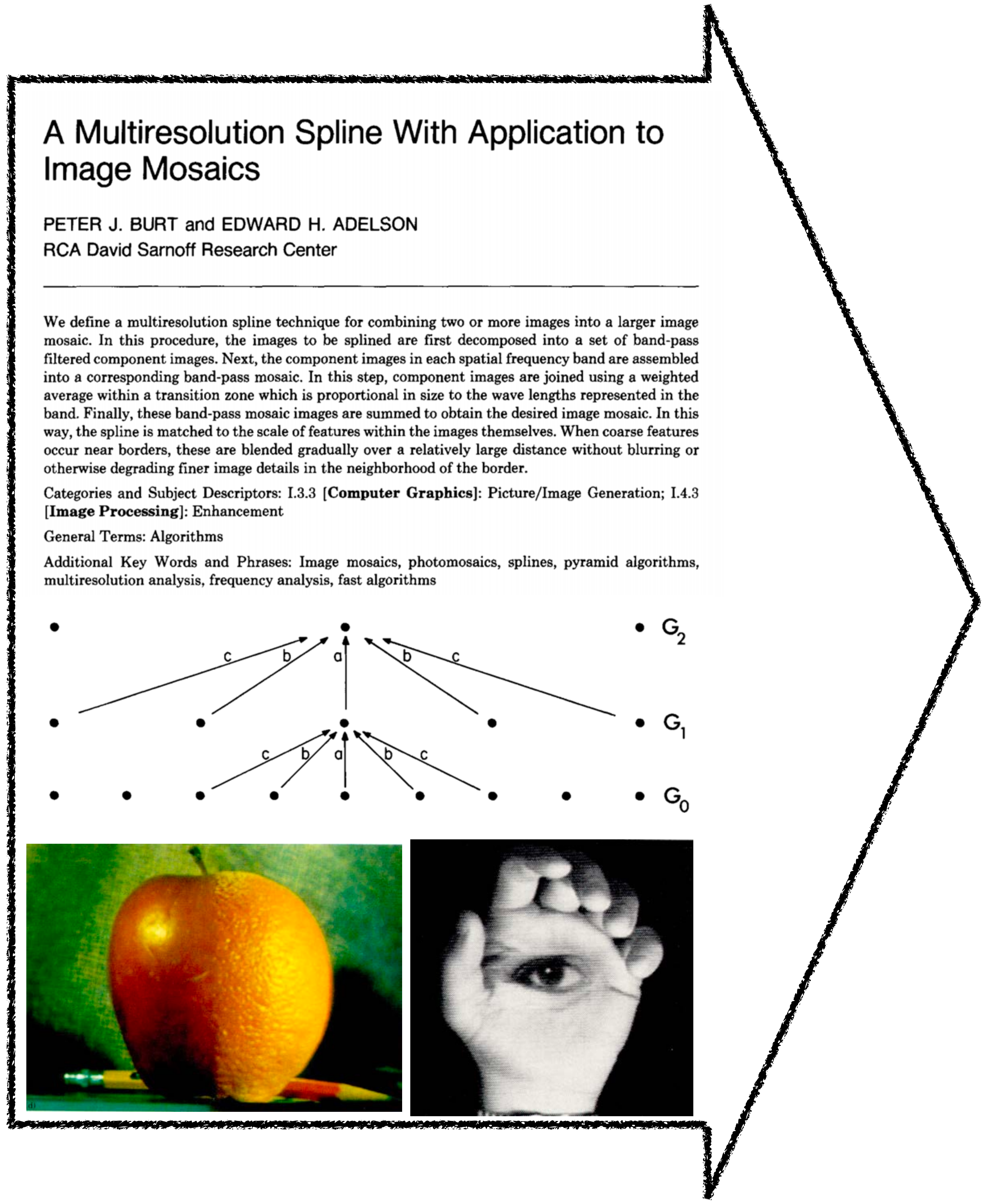






Fast local Laplacian

Basis for Adobe Photoshop & Lightroom tone adjustments



A Multiresolution Spline With Application to Image Mosaics

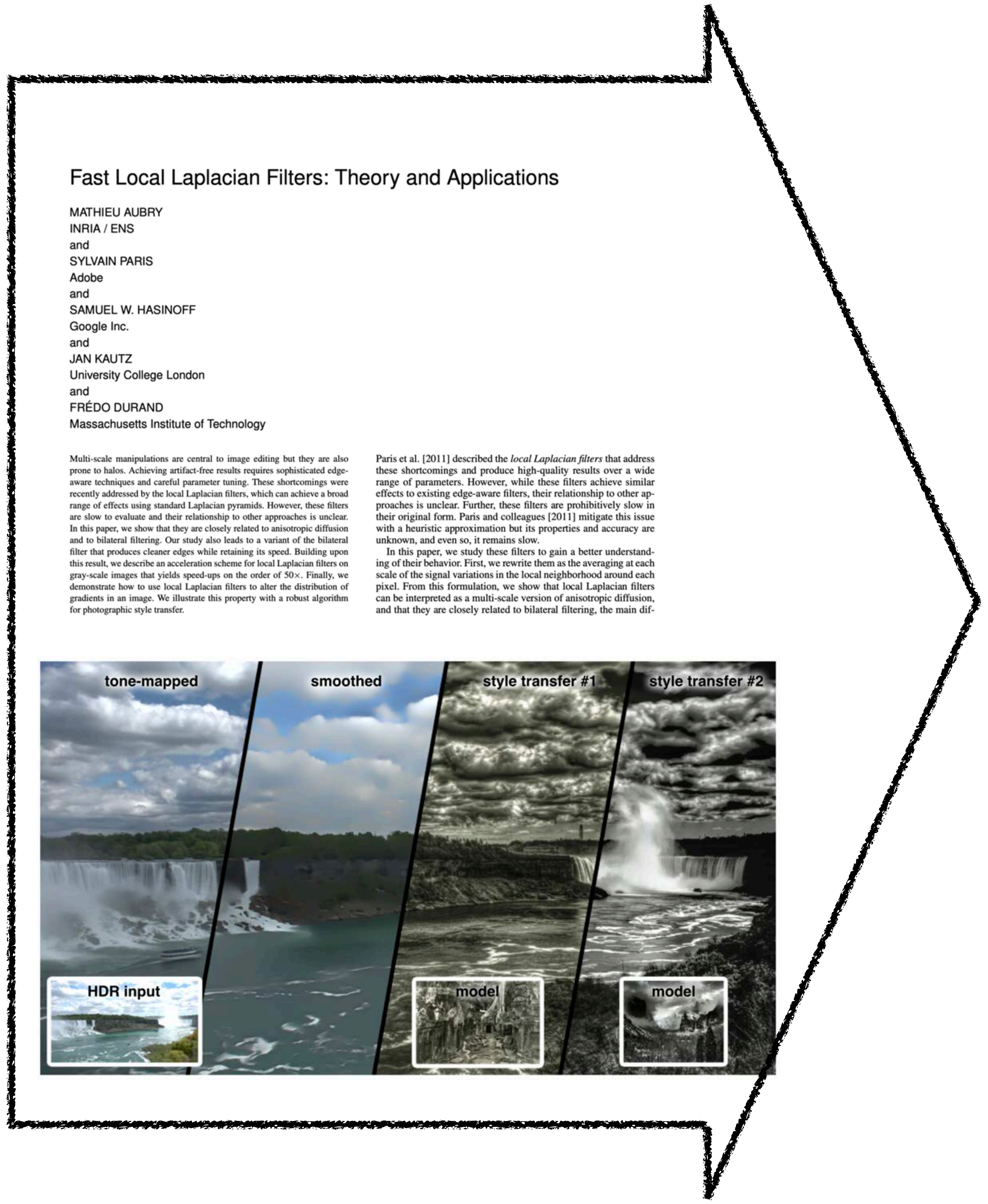
PETER J. BURT and EDWARD H. ADELSON
RCA David Sarnoff Research Center

We define a multiresolution spline technique for combining two or more images into a larger image mosaic. In this procedure, the images to be splined are first decomposed into a set of band-pass filtered component images. Next, the component images in each spatial frequency band are assembled into a corresponding band-pass mosaic. In this step, component images are joined using a weighted average within a transition zone which is proportional in size to the wave lengths represented in the band. Finally, these band-pass mosaic images are summed to obtain the desired image mosaic. In this way, the spline is matched to the scale of features within the images themselves. When coarse features occur near borders, these are blended gradually over a relatively large distance without blurring or otherwise degrading finer image details in the neighborhood of the border.

Categories and Subject Descriptors: I.3.3 [**Computer Graphics**]: Picture/Image Generation; I.4.3 [**Image Processing**]: Enhancement

General Terms: Algorithms

Additional Key Words and Phrases: Image mosaics, photomosaics, splines, pyramid algorithms, multiresolution analysis, frequency analysis, fast algorithms



Fast Local Laplacian Filters: Theory and Applications

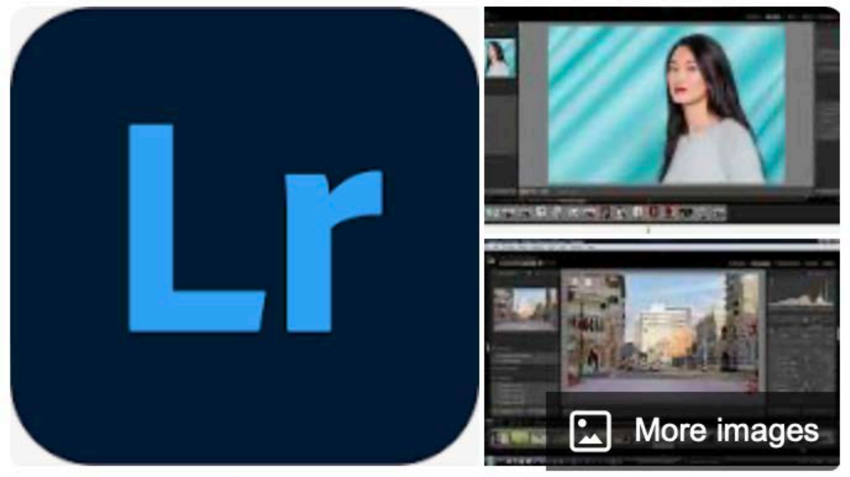
MATHIEU AUBRY
INRIA / ENS
and
SYLVAIN PARIS
Adobe
and
SAMUEL W. HASINOFF
Google Inc.
and
JAN KAUTZ
University College London
and
FRÉDO DURAND
Massachusetts Institute of Technology

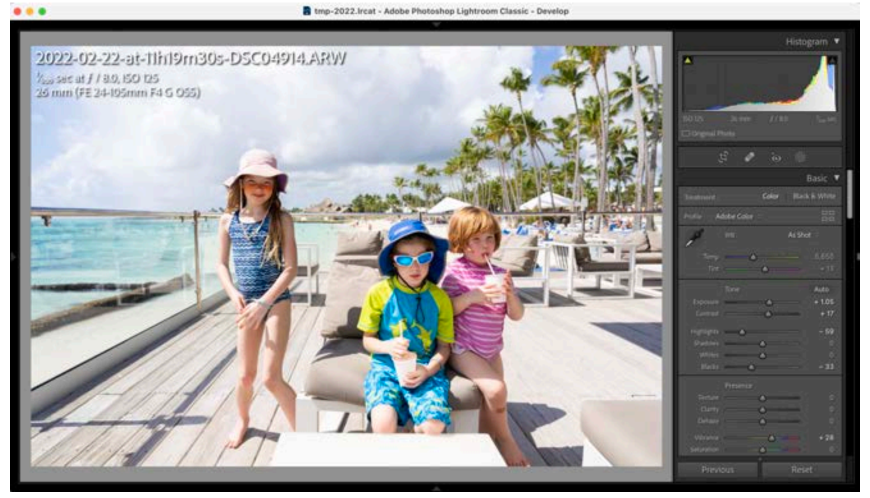
Multi-scale manipulations are central to image editing but they are also prone to halos. Achieving artifact-free results requires sophisticated edge-aware techniques and careful parameter tuning. These shortcomings were recently addressed by the local Laplacian filters, which can achieve a broad range of effects using standard Laplacian pyramids. However, these filters are slow to evaluate and their relationship to other approaches is unclear. In this paper, we show that they are closely related to anisotropic diffusion and to bilateral filtering. Our study also leads to a variant of the bilateral filter that produces cleaner edges while retaining its speed. Building upon this result, we describe an acceleration scheme for local Laplacian filters on gray-scale images that yields speed-ups on the order of 50×. Finally, we demonstrate how to use local Laplacian filters to alter the distribution of gradients in an image. We illustrate this property with a robust algorithm for photographic style transfer.

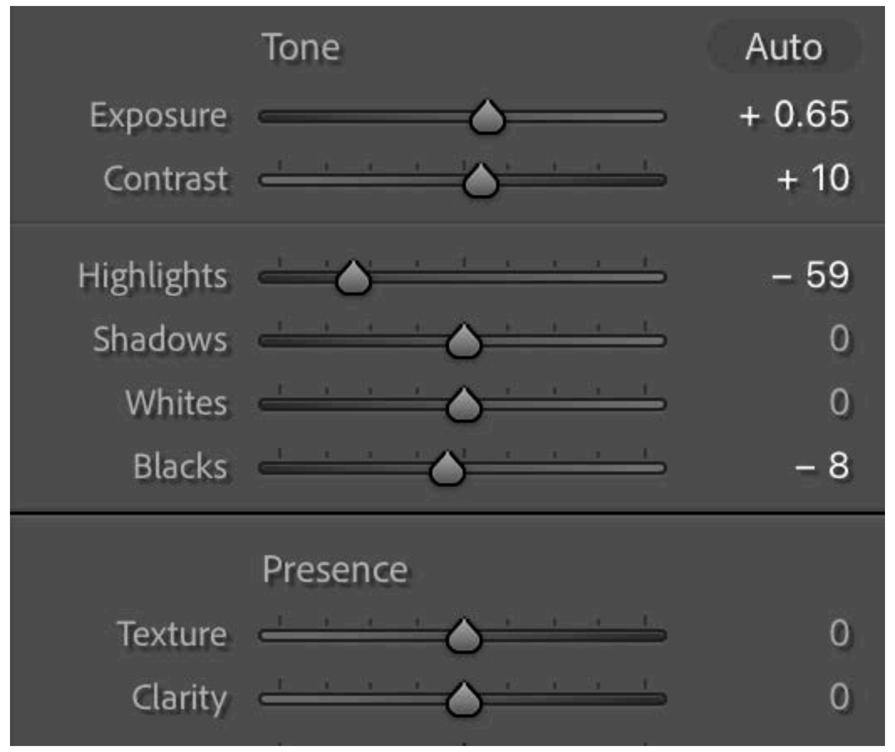
Paris et al. [2011] described the *local Laplacian filters* that address these shortcomings and produce high-quality results over a wide range of parameters. However, while these filters achieve similar effects to existing edge-aware filters, their relationship to other approaches is unclear. Further, these filters are prohibitively slow in their original form. Paris and colleagues [2011] mitigate this issue with a heuristic approximation but its properties and accuracy are unknown, and even so, it remains slow.

In this paper, we study these filters to gain a better understanding of their behavior. First, we rewrite them as the averaging at each scale of the signal variations in the local neighborhood around each pixel. From this formulation, we show that local Laplacian filters can be interpreted as a multi-scale version of anisotropic diffusion, and that they are closely related to bilateral filtering, the main dif-

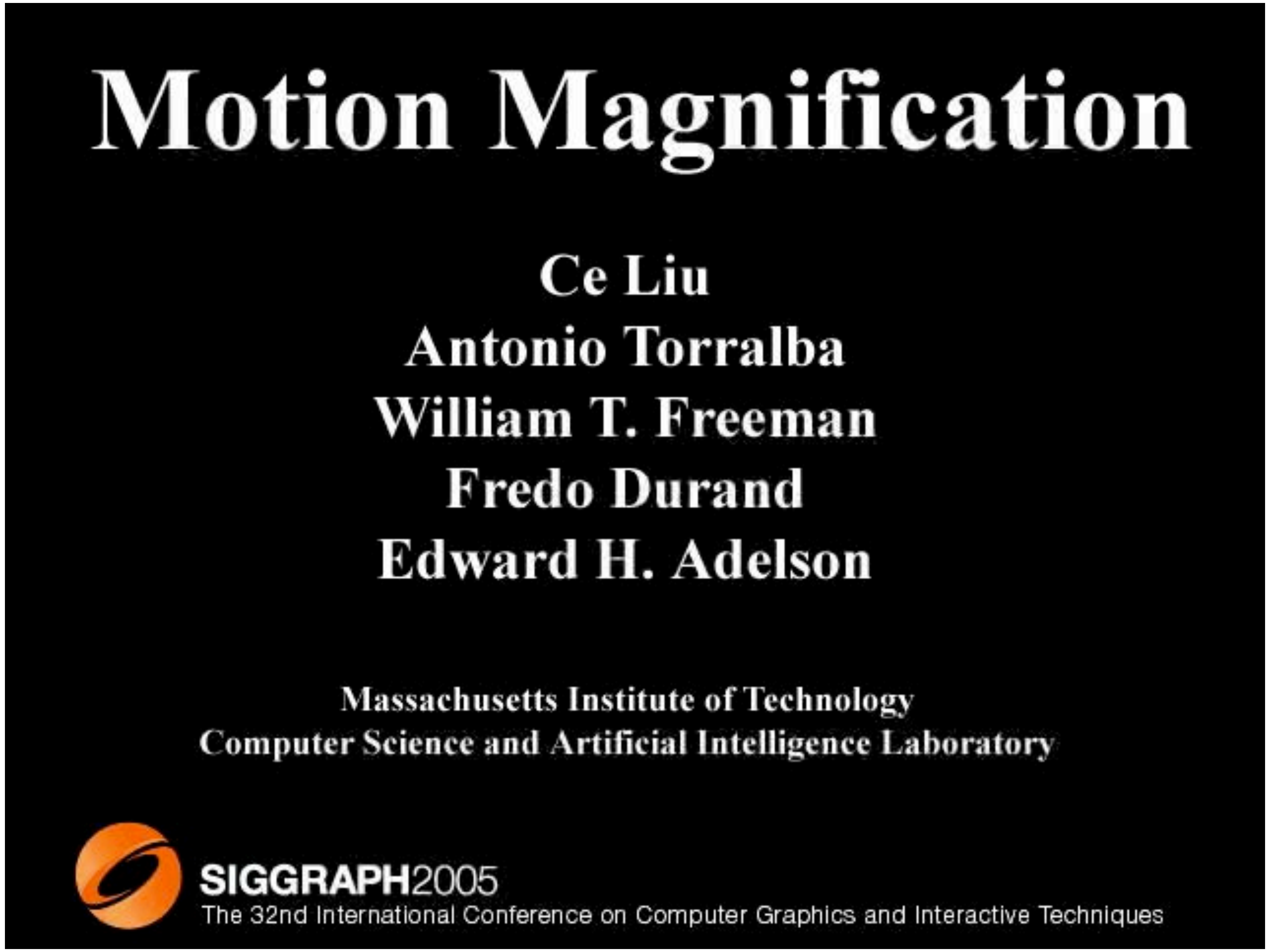
Adobe Lightroom
Software







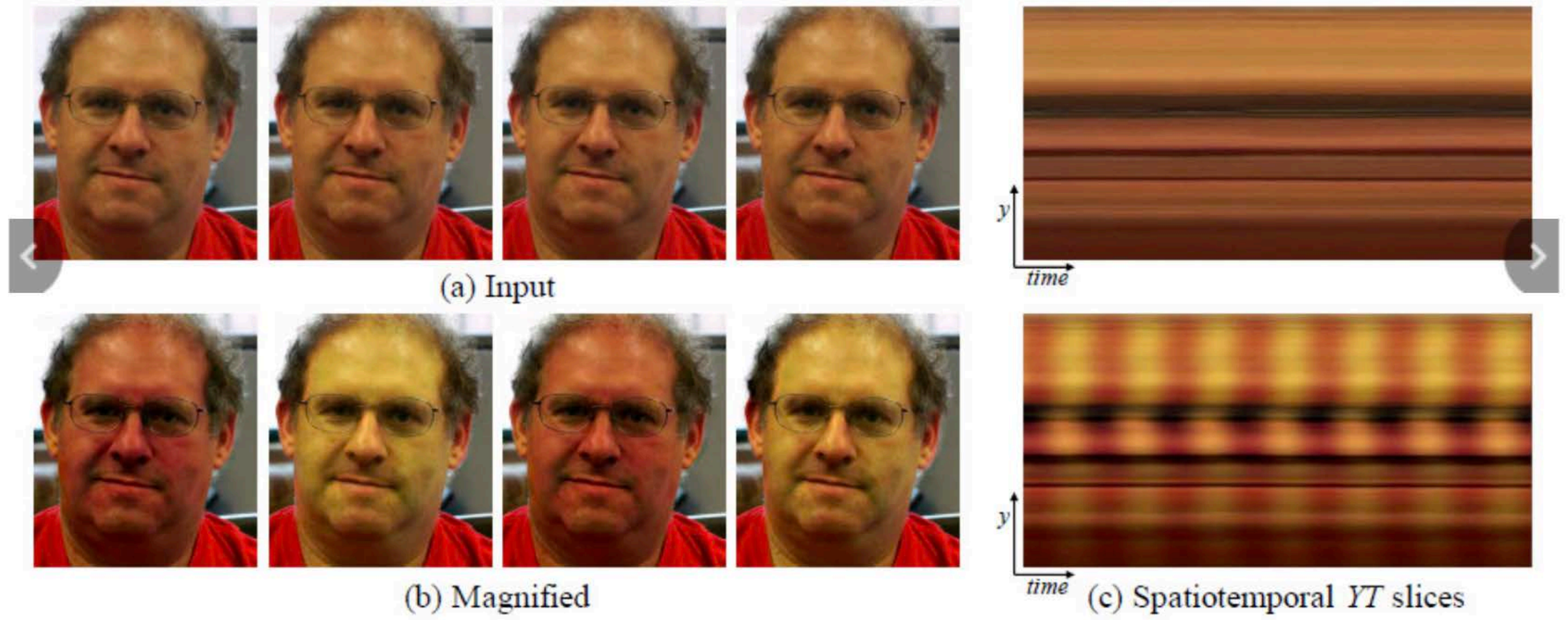
Motion magnification



Video Magnification

CSAIL

Videos Software Publications Applications People Related Work Talks



An example of using our Eulerian Video Magnification framework for visualizing the human pulse. (a) Four frames from the original video sequence. (b) The same four frames with the subject's pulse signal amplified. (c) A vertical scan line from the input (top) and output (bottom) videos plotted over time shows how our method amplifies the periodic color variation. In the input sequence the signal is imperceptible, but in the magnified sequence the variation is clear.

Many seemingly static scenes contain subtle changes that are invisible to the naked human eye. However, it is possible to pull out these small changes from videos through the use of algorithms we have developed. We give a way to visualize these small changes by amplifying them and we present algorithms to pull out interesting signals from these videos, such as the human pulse, sound from vibrating objects and the motion of hot air.

Like 3.6K Share

A lot of my graphics and computational photography work has to do with the Plenoptic function

and the insight that light field, space time, etc. are all similar

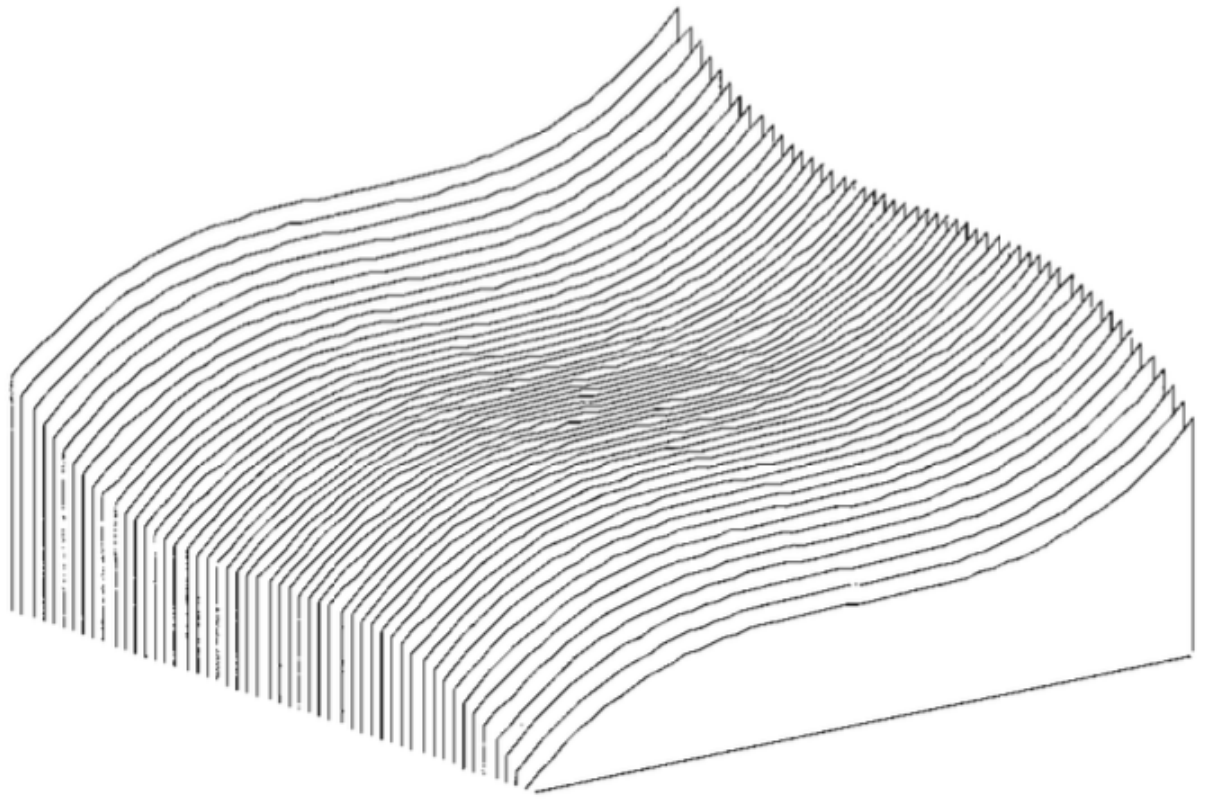
Wavefront coding in the space of light rays

New paradigm for imaging systems

W. Thomas Cathey and Edward R. Dowski

We describe a new paradigm for designing hybrid imaging systems. These imaging systems use optics with a special aspheric surface to code the image so that the point-spread function or the modulation transfer function has specified characteristics. Signal processing then decodes the detected image. The coding can be done so that the depth of focus can be extended. This allows the manufacturing tolerance to be reduced, focus-related aberrations to be controlled, and imaging systems to be constructed with only one optical element plus some signal processing.

OCIS codes: 080.3620, 110.0110, 110.2990, 110.0180, 110.4850, 180.0180.



Plenoptic function insight

The Plenoptic Function and the Elements of Early Vision

Edward H. Adelson and James R. Bergen

What are the elements of early vision? This question might be taken to mean, What are the fundamental atoms of vision?—and might be variously answered in terms of such candidate structures as edges, peaks, corners, and so on. In this chapter we adopt a rather different point of view and ask the question, What are the fundamental *substances* of vision? This distinction is important because we wish to focus on the first steps in extraction of visual information. At this level it is premature to talk about discrete objects, even such simple ones as edges and corners.

$$P = P(\theta,\phi,\lambda,t,V_x,V_y,V_z).$$

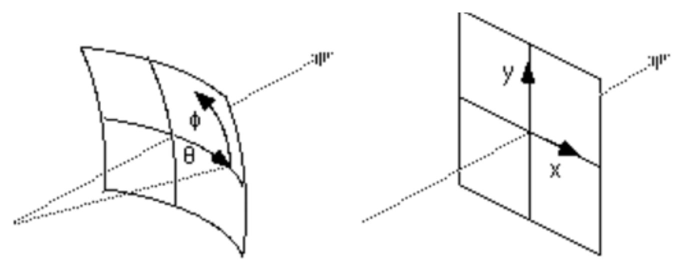
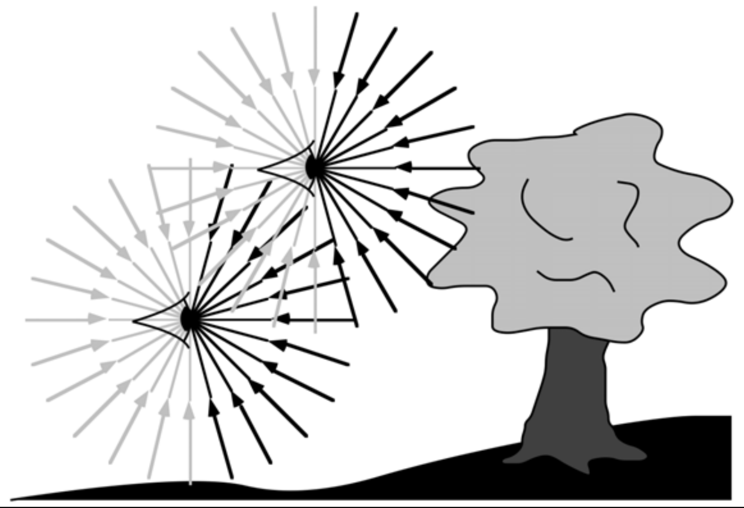


Fig.1.2
The image information available from a single viewing position is defined by the pencil of light rays passing through the pupil. The rays may be parameterized in angular coordinates or in Cartesian coordinates. The Cartesian approach is commonly used in machine vision and computer graphics, but the angular approach can more easily represent the full sphere of

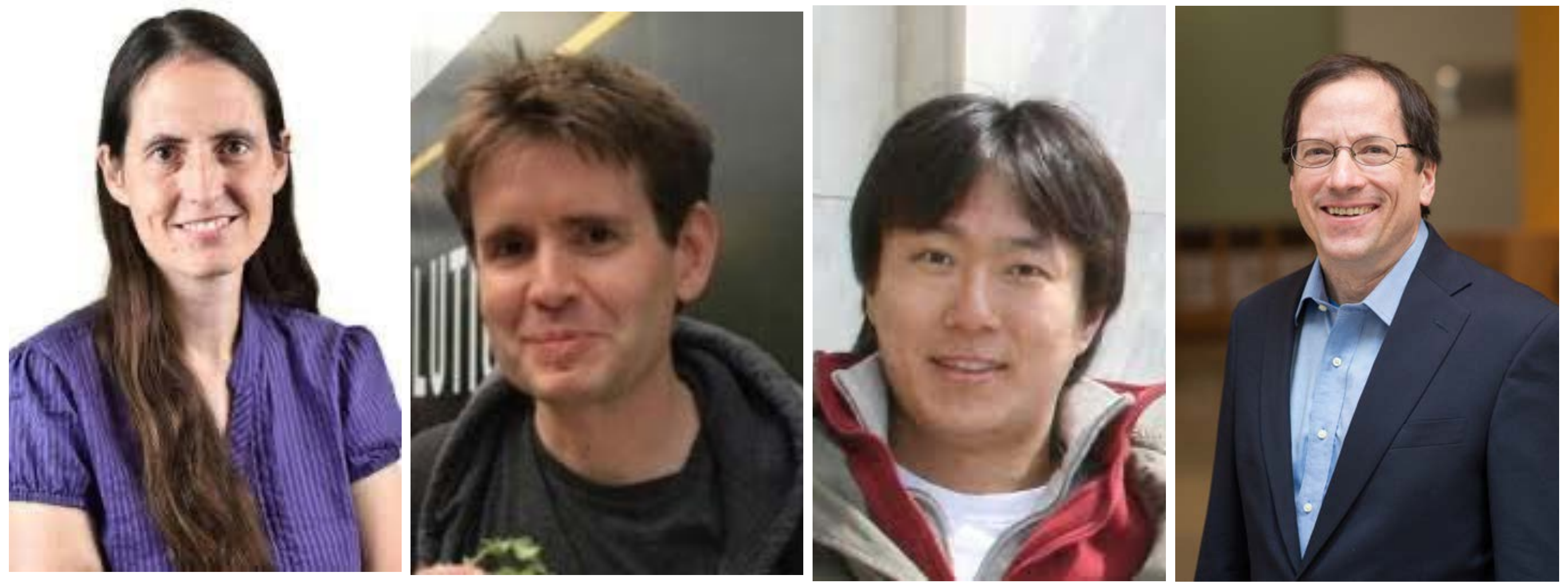


Same coding but in space-time

Motion-Invariant Photography

Anat Levin Peter Sand Taeg Sang Cho Frédo Durand William T. Freeman
Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory

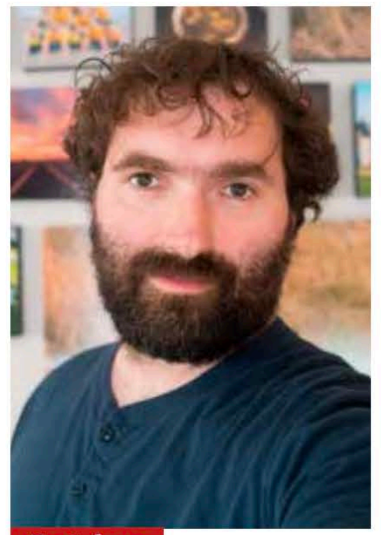
Figure 1: Left: *Blurred motion captured by a static camera.* **Center:** *The same scene captured by a camera with a specially designed motion that causes both the static and dynamic regions to blur identically.* **Right:** *The blur from the center image can be removed independently of motion via deconvolution of the entire image with a single known point spread function.*



I was recently asked what the biggest photography innovations were since the advent of digital

RÉPONSES GRAND FORMAT

FRÉDO DURAND



En 6 dates

- ➜ **1973 :** Naissance à L'Haÿ-les-Roses
- ➜ **1981 :** Premier appareil photo offert par son père, un Kodak Retinette
- ➜ **1993 :** Entre à l'ENS rue d'Ulm
- ➜ **1999 :** Doctorat à l'INRIA Grenoble avec Claude Puech et George Drettakis
- ➜ **2002 :** Professeur au Massachusetts Institute of Technology (États-Unis)

Peu de photographes ont entendu prononcer son nom, ils sont pourtant des millions à utiliser quotidiennement le fruit de ses travaux. Frédo Durand, professeur au MIT (Massachusetts Institute of Technology à Cambridge, États-Unis) est depuis 20 ans au centre des multiples développements que l'on trouve au cœur de Photoshop, Lightroom, Camera Raw, mais aussi dans le firmware de nos appareils photos, dans les systèmes embarqués et les apps de nos smartphones. À la fois acteur et témoin privilégié de l'évolution des technologies photographiques, ce Français de 49 ans – pour ne rien gâter photographe fervent et lecteur de Réponses Photo – nous est donc apparu comme l'interlocuteur idéal pour jauger le chemin parcouru ces trente dernières années, depuis la naissance de notre magazine, et pour nous projeter dans l'avenir de la photographie, tant d'un point de vue technique que pratique. **Propos recueillis par Yann Garret**

Comment votre parcours scientifique a-t-il croisé la photographie ?
J'ai toujours été fasciné par l'image et j'ai fait mon doctorat sur l'image de synthèse 3D et la simulation photoréaliste de l'éclairage à l'université de Grenoble. Ces travaux sur les images virtuelles m'ont naturellement amené à me poser des questions sur les images que prennent les photographes dans le monde réel et des questions sur la perception visuelle. J'ai décidé de créer un cours, "The art and science of depiction"(1) pour explorer l'interaction entre synthèse d'images, arts visuels (dont la photo) et perception humaine.

Il y a 30 ans, quel était l'état de l'art en matière de technologies photographiques ?
À l'époque, l'autofocus commençait à dominer les reflex, et la stabilisation d'image et la fluorite révolutionnaient les téléobjectifs. Les objectifs zoom avaient encore mauvaise réputation chez les professionnels. La photographie de qualité restait technique, même si les automatismes tels que l'autofocus et l'exposition commençaient a grandement aider les amateurs. Kodak venait de sortir le premier reflex numérique commercial, le DCS 100, une énorme brique avec une définition de 1,3 mégapixel qui coûtait 30000$ tout de même, destiné à accélérer la transmission des photos pour les photojournalistes. À ce prix-là, même pas d'écran !
Au début des années 1990, les ordinateurs accéléraient encore de façon continue, suivant la fameuse loi de Moore, et leur capacité de calcul doublait tous les deux ans. Les batterie ion-lithium venaient de sortir (Sony, 1991), et la plupart des écrans étaient encore de gros CRT basse résolution avec une densité de 72 dpi. L'ordinateur portable Macintosh PowerBook avait un processeur à 25 MHz, 8 Mo de RAM, jusqu'à 80 Mo de disque dur (le poids d'une image RAW moderne !) et un écran LCD monochrome de 640x400. Les premières imprimantes à jet d'encre couleur 300 dpi venait de sortir. Elles utilisaient 4 encres (CMYK) et étaient loin d'offrir une qualité photographique.

Et puis vient le tournant de l'an 2000…
Oui, c'est à ce moment-là que la réalité commerciale des appareils numériques se concrétise. Le Nikon D1, en1999, a été le premier appareil numérique à connaître un véritable usage chez les pros. J'ai eu la chance de travailler avec un D1 et c'est vraiment l'appareil qui m'a donné le goût de la photo numérique. Une mention aussi pour le Canon D30 en 2000, qui a lancé les reflex numériques pour le plus grand public, malgré un prix assez élevé.
Et puis les premiers téléphones portables avec une caméra ont été lancés en 2000 par Samsung et Sharp (0,1 mégapixel seulement !). Mais je dirais que c'est avec le Nokia N95 que la qualité de la caméra (5 MP) et de l'écran ainsi que l'accélération des réseaux avec la 3,5G, ont vraiment rendu les téléphones portables attractifs pour la photo. Je ne surprendrai personne aujourd'hui en disant que l'entrée des portables dans le monde de la photographie est l'événement le plus significatif depuis l'invention du numérique. L'industrie ne sera plus jamais la même. Les téléphones portables ont aussi accéléré la révolution computationnelle car leurs contraintes physiques rendent la création d'optique et de capteurs de qualité à des coûts acceptables très difficile.
Il est aussi intéressant de noter qu'il y a trente ans, le monde de la vidéo et de la photo étaient complètement séparés et demandaient des appareils différents. En 2001, le compact Canon Pro 90 IS fut, je crois, le premier à offrir de la vidéo en plus des photos. Les reflex furent longs à suivre car les capteurs chauffaient. C'est en 2008 que le D90 et surtout le Canon 5D Mark II ont permis de capturer des vidéos HD et commencé à révolutionner le monde de la vidéo, en proposant une qualité d'image qui jusque-là demandait un équipement qui coûtait plus de cent mille euros. C'est que les capteurs de ces reflex étaient bien plus gros que ceux des caméscopes même haut de gamme, ce qui permettait des images en basse lumière et des effets de profondeur de champ jusque-là impossibles.

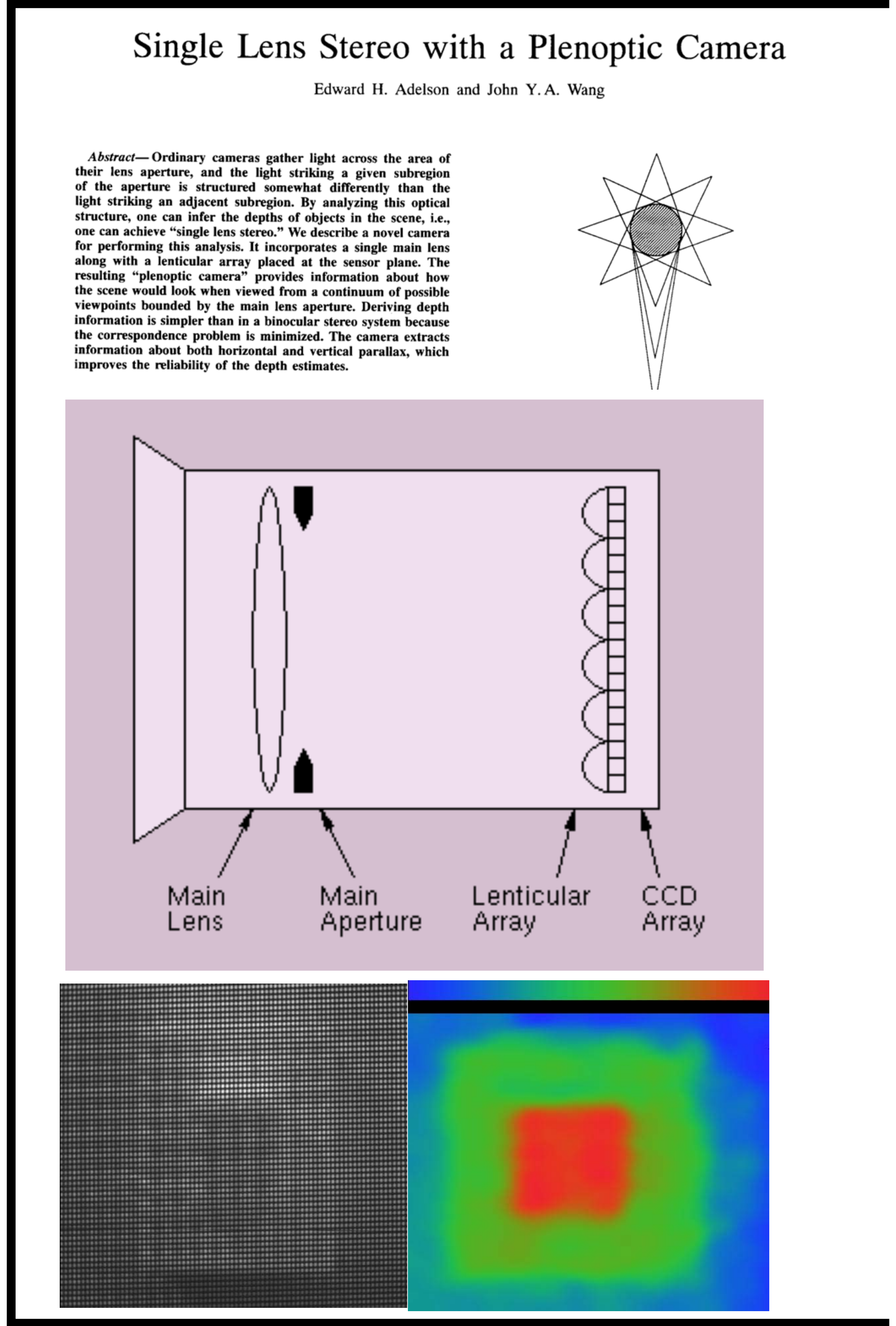
Depuis lors, peut-on parler de progrès continus ?
Le Nikon D3s ,en 2009, a marqué pour moi la fin, ou tout du moins le gros ralentissement, de l'amélioration des capteurs en

(1) *https://people.csail.mit.edu/fredo/ArtAndScienceOfDepiction/*

Titre légende
Tem aperorum quaeprae nobis doluptaque invellanRia aut aut aut fuga. Et fugia sinctem natem eos doluptaturit qui doloribus nonsequi volor abore non reptatur, qui utatest videstrumBusant que iusam iligent utenihi tiores sus mi, aut ut venda dem ut prehent

Plenoptic/dual-pixel on-sensor autofocus

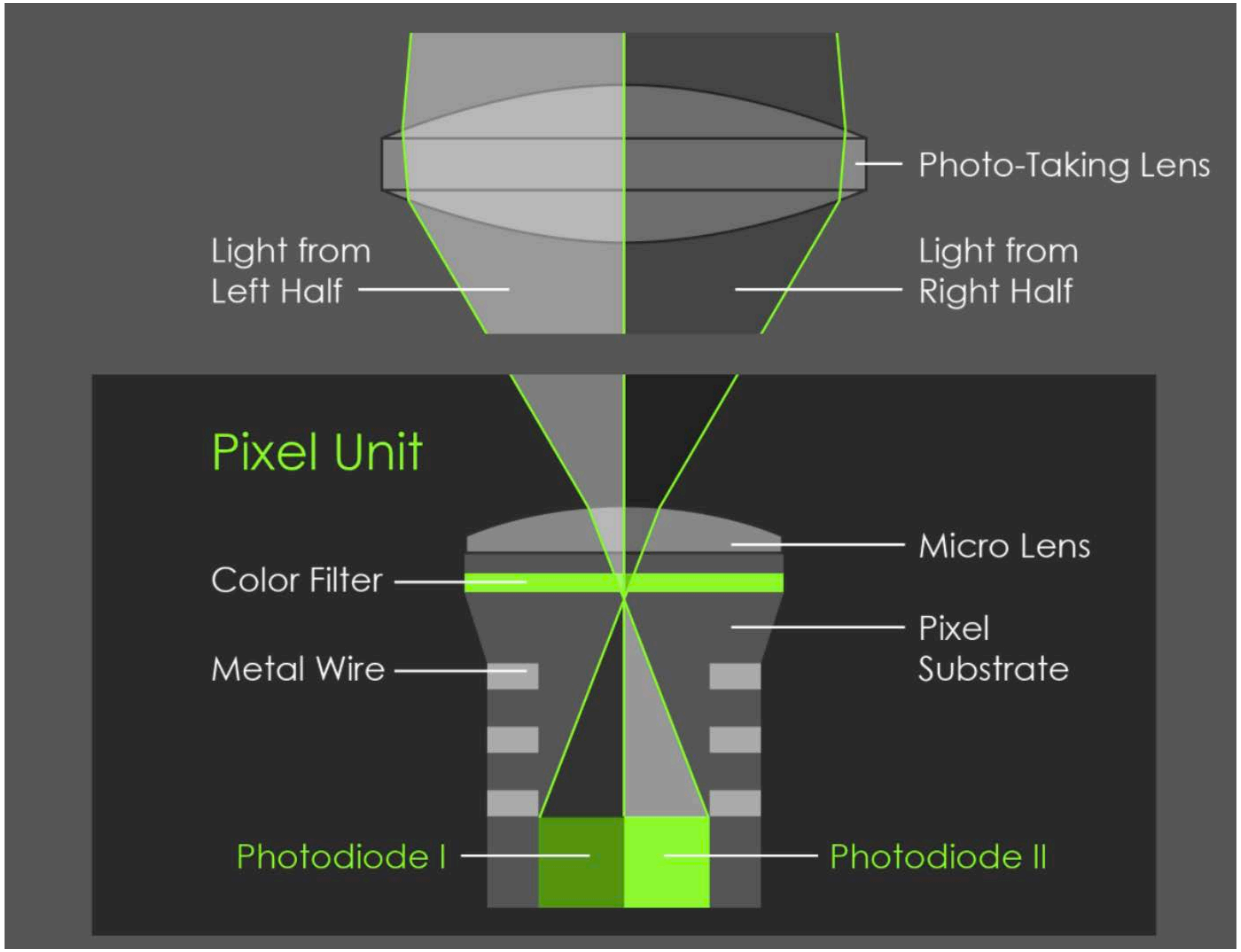
Biggest innovation since the beginning of digital photography



Single Lens Stereo with a Plenoptic Camera

Edward H. Adelson and John Y. A. Wang

Abstract— Ordinary cameras gather light across the area of their lens aperture, and the light striking a given subregion of the aperture is structured somewhat differently than the light striking an adjacent subregion. By analyzing this optical structure, one can infer the depths of objects in the scene, i.e., one can achieve "single lens stereo." We describe a novel camera for performing this analysis. It incorporates a single main lens along with a lenticular array placed at the sensor plane. The resulting "plenoptic camera" provides information about how the scene would look when viewed from a continuum of possible viewpoints bounded by the main lens aperture. Deriving depth information is simpler than in a binocular stereo system because the correspondence problem is minimized. The camera extracts information about both horizontal and vertical parallax, which improves the reliability of the depth estimates.

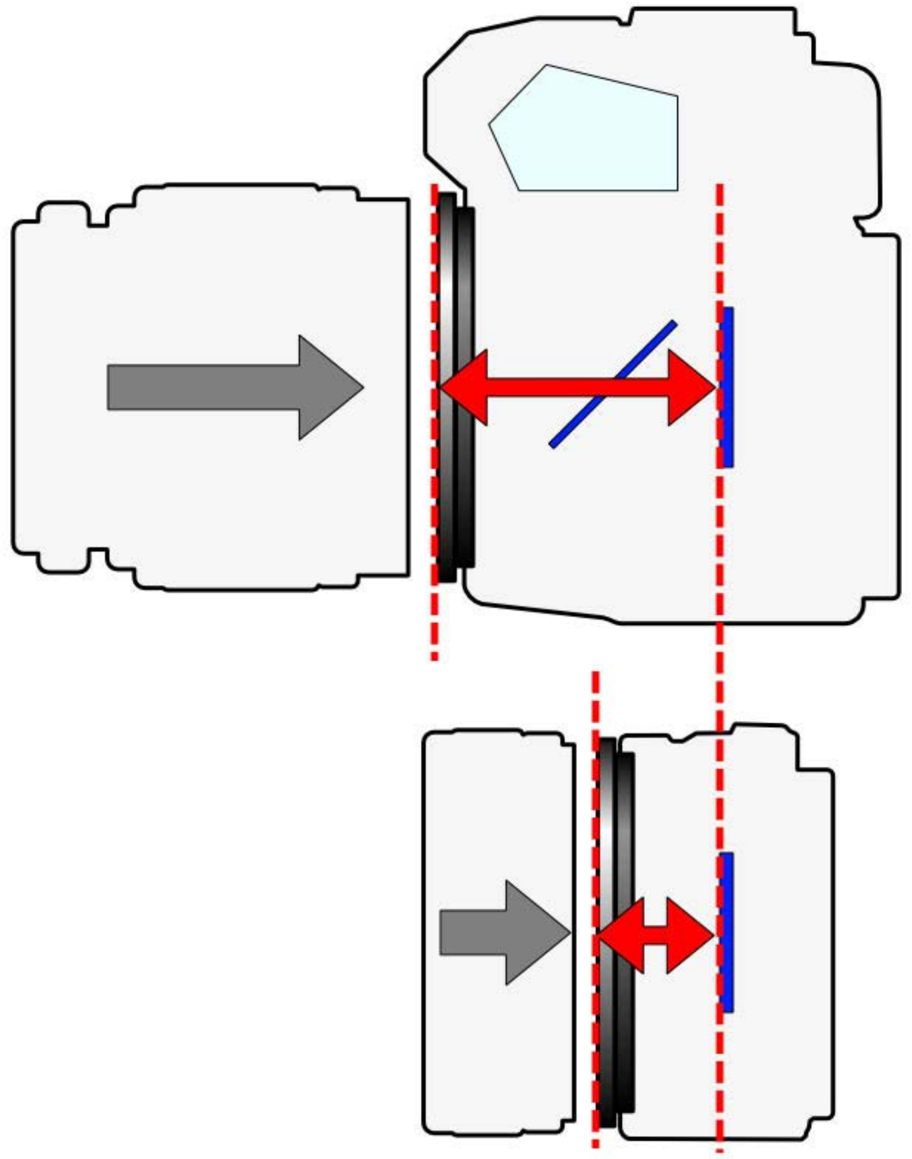


Impact of plenoptic/dual-pixel on-sensor AF

Enables phase-based (~stereo) autofocus on sensor, without mirror

Made mirrorless competitive with DSLR

- Can focus while recording a Video
- Can perform face/eye detection for AF because you have access to the sensor
- Makes mechanical design much simpler (no need for mirror, etc.) -> cheaper, more reliable
- Removing the mirror makes lens design easier (more space) (e.g. Canon 24-70 f/2 was impossible before.)
- https://www.diyphotography.net/the-dslr-will-probably-die-are-mirrorless-the-future-of-large-standalone-cameras/



Apparent Ridges

Started with an insight from Ted

- First, human perception is sensitive to the variation of shading, and since shape perception is little affected by lighting and reflectance modification, we should focus on normal variation.
- Second, view-dependent lines better convey smooth surfaces.
- From this we define view-dependent curvature, and apparent ridges as the loci of points that maximize a view-dependent curvature.

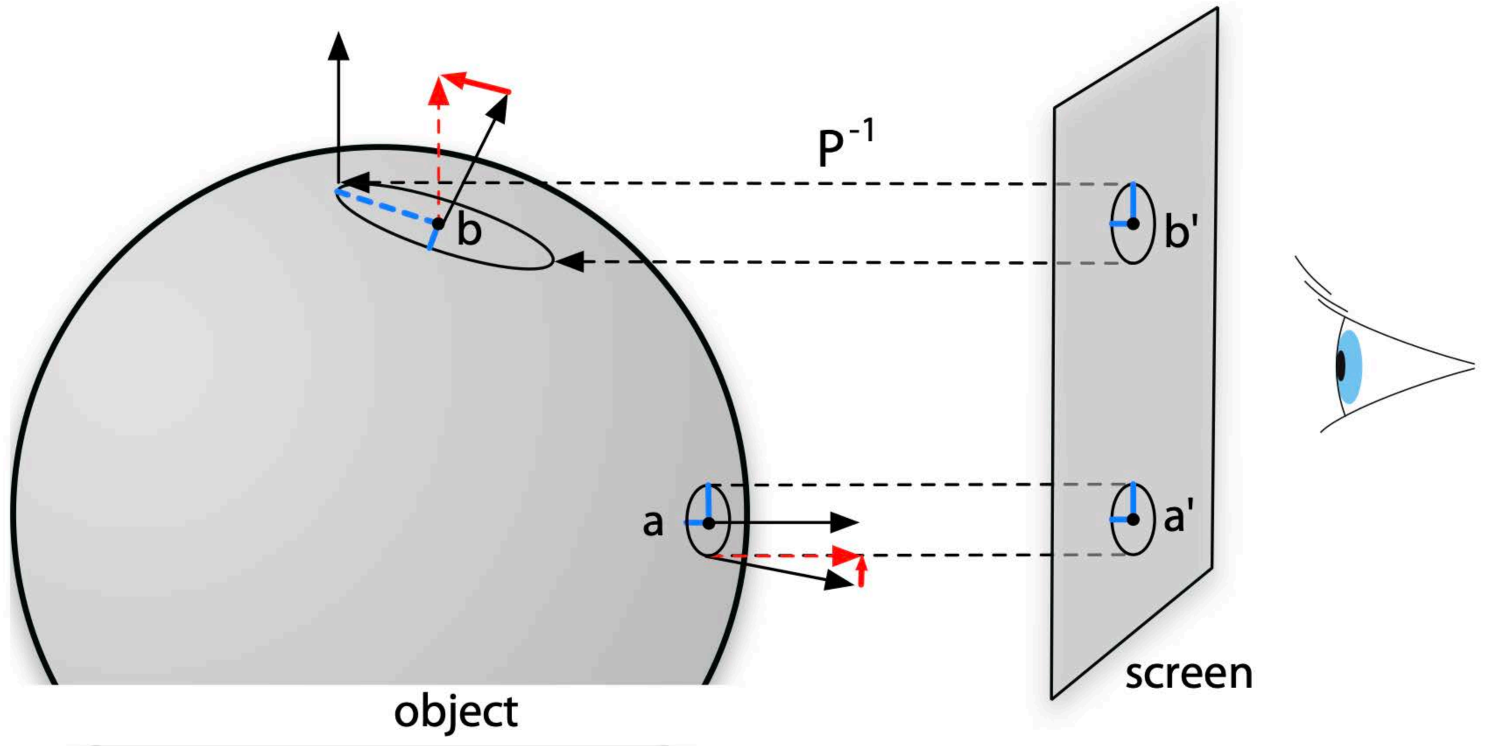


Figure 4: *The maximum view-dependent curvature at b' is much larger than at a' uniquely because of projection.*

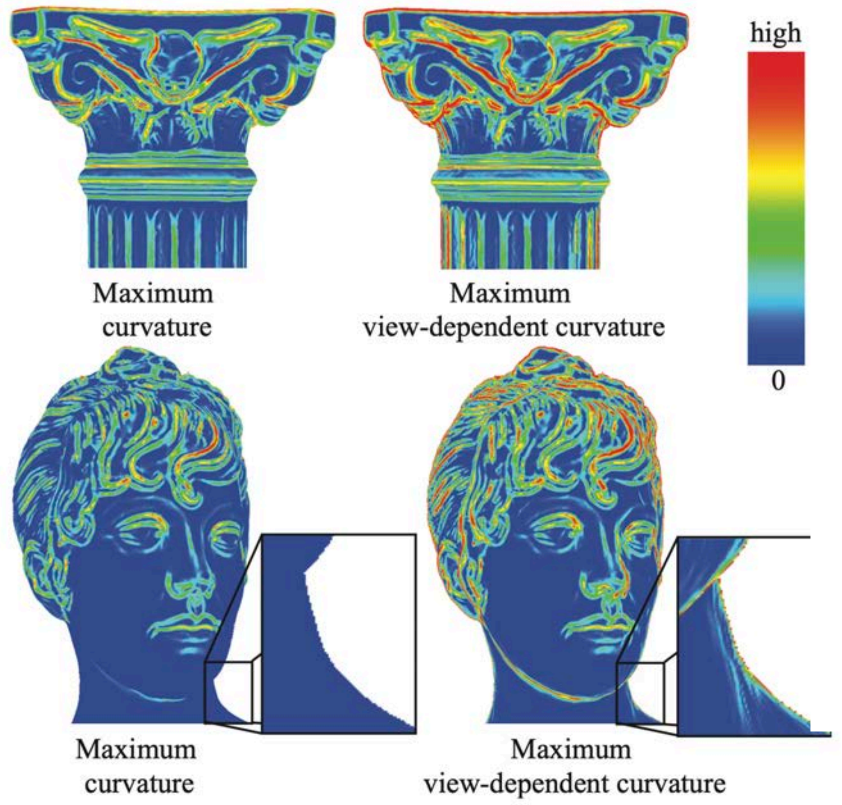
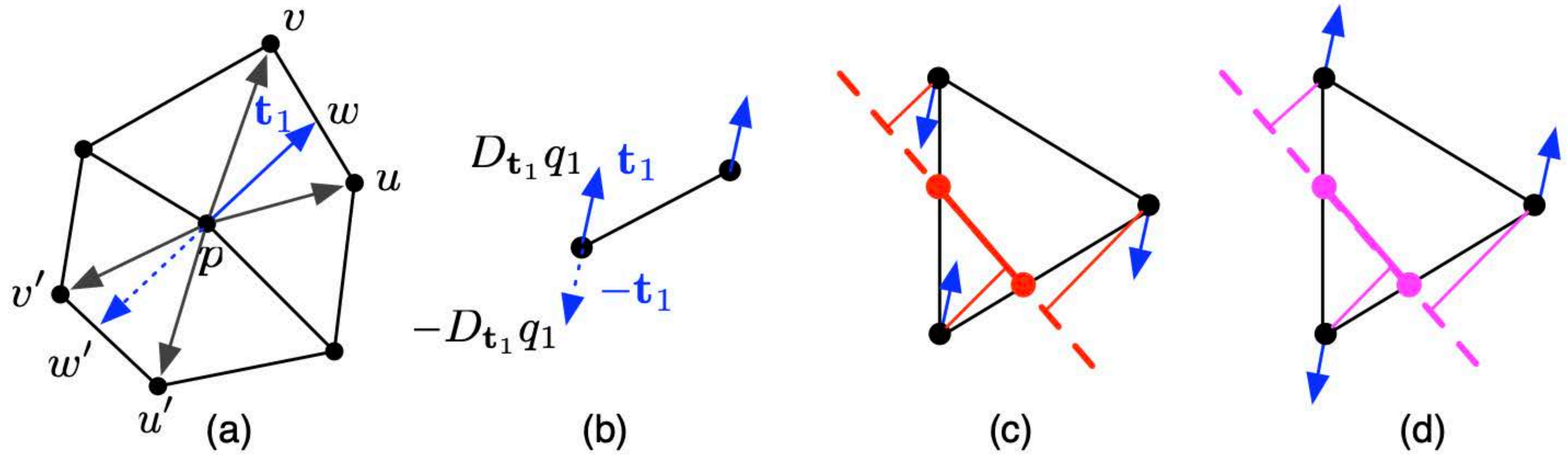


Figure 5: *Comparison of curvature and view-dependent curvature. At front facing parts of the object, the values are similar. As the object normal turns away from the viewer, view-dependent curvature becomes much larger due to projection. View-dependent curvature approaches a maximum of infinity at the contours and so contours*



Transactions on Graphics, Proceedings of SIGGRAPH 2007

Apparent Ridges

Apparent Ridges for Line Drawing

Tilke Judd[1] Frédo Durand[1] Edward Adelson[1,2]

[1]MIT Computer Science and Artificial Intelligence Laboratory [2] MIT Dept. of Brain and Cognitive Sciences

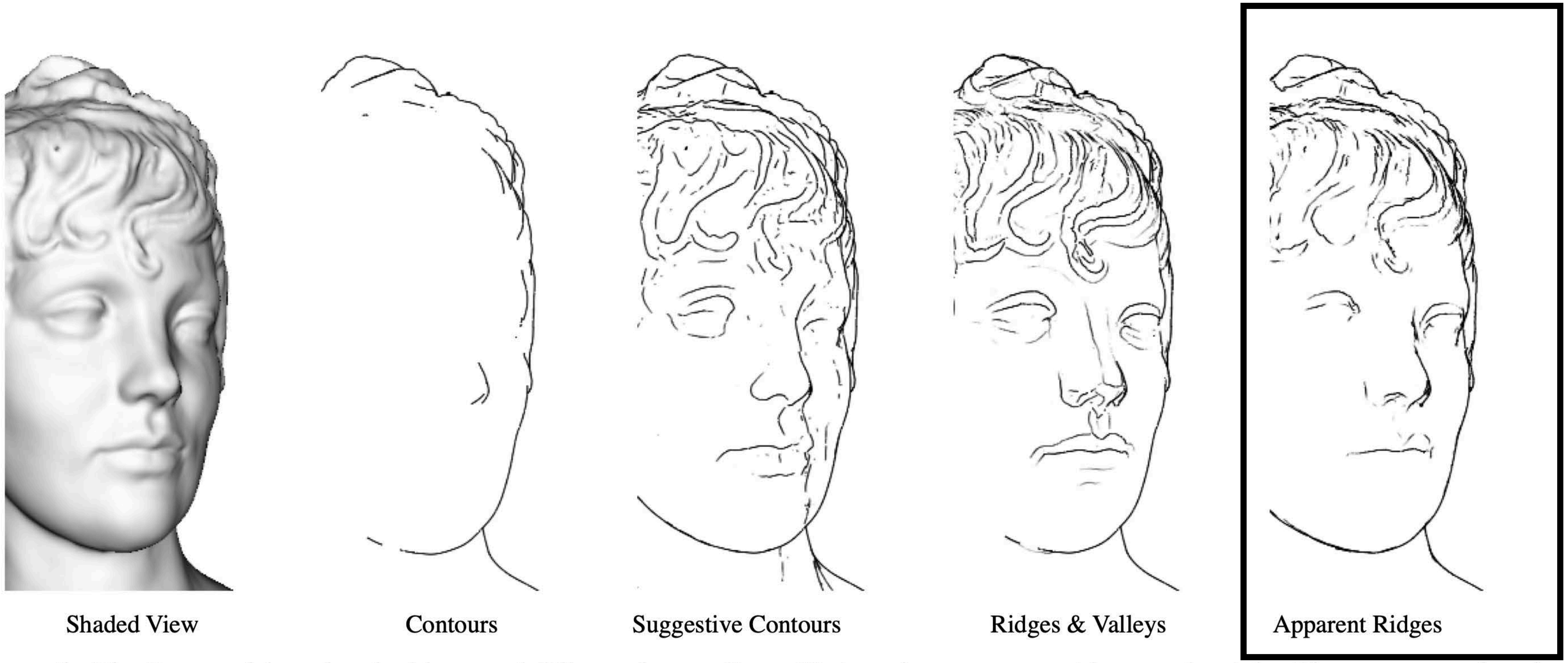


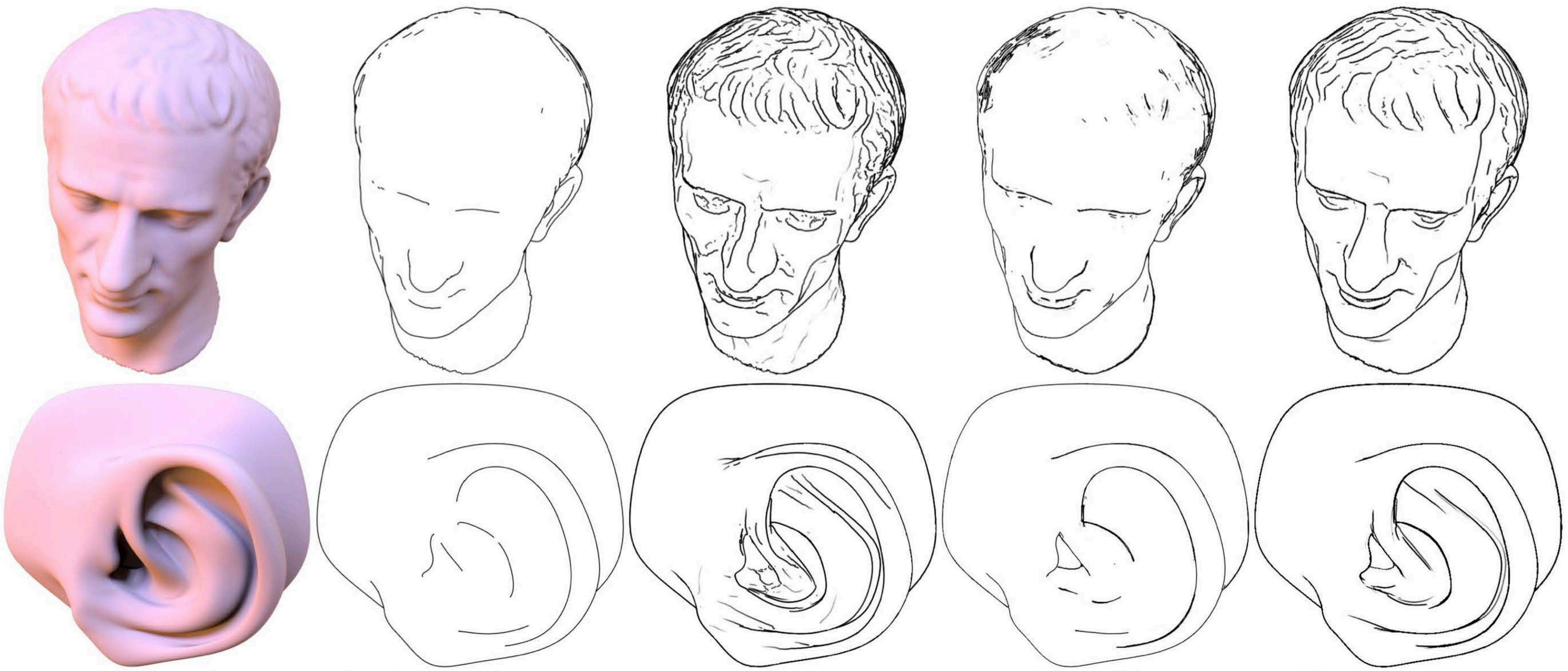
Figure 1: *The Bust model rendered with several different feature lines. We introduce apparent ridges on the right. They correspond to the maxima of the normal variation with respect to the viewing plane. Note in particular the left side of the face (to the right) in the suggestive contour drawing and the nose drawn with ridges and valleys.*

15 years later, still competitive

Sure, state of the art with deep learning is better

Neural Contours: Learning to Draw Lines from 3D Shapes

Difan Liu[1] Mohamed Nabail[1] Aaron Hertzmann[2] Evangelos Kalogerakis[1]
[1]University of Massachusetts Amherst [2]Adobe Research



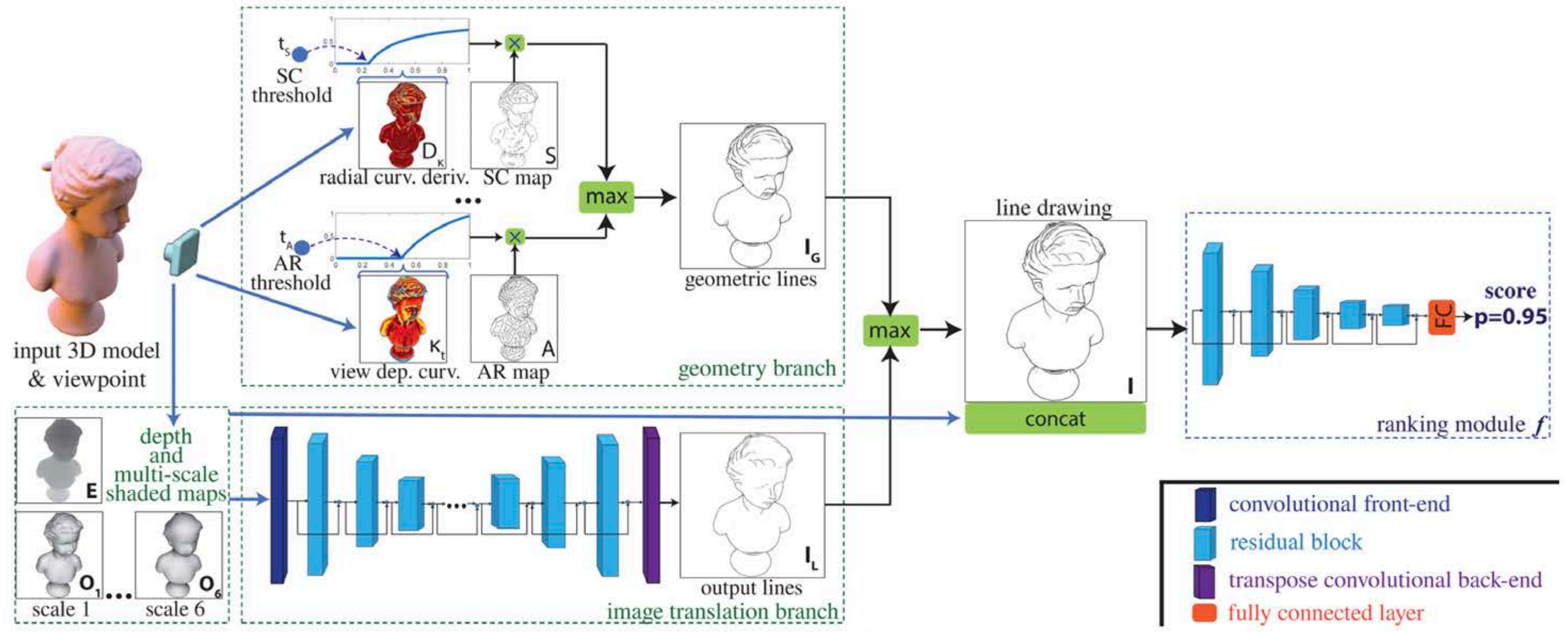


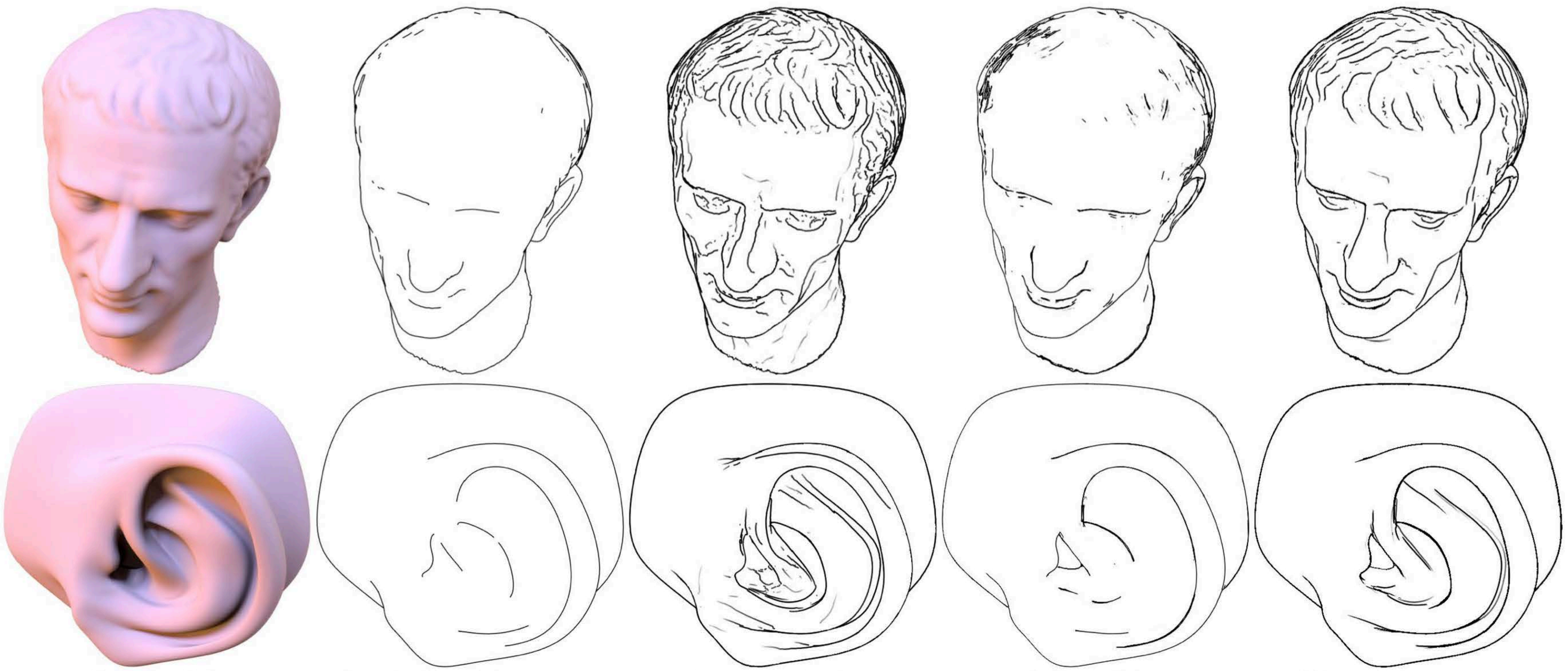
Figure 3: Our network architecture: the input 3D model is processed by a geometry branch operating on curvature features, and an image-based branch operating on view-based representations. Their outputs are combined to create a line drawing, which is in turn evaluated by a ranking module that helps determining optimal line drawing parameters.

15 years later, still competitive

State of the art with deep learning is better, but still relies on apparent ridges

Neural Contours: Learning to Draw Lines from 3D Shapes

Difan Liu[1] Mohamed Nabail[1] Aaron Hertzmann[2] Evangelos Kalogerakis[1]
[1]University of Massachusetts Amherst [2]Adobe Research



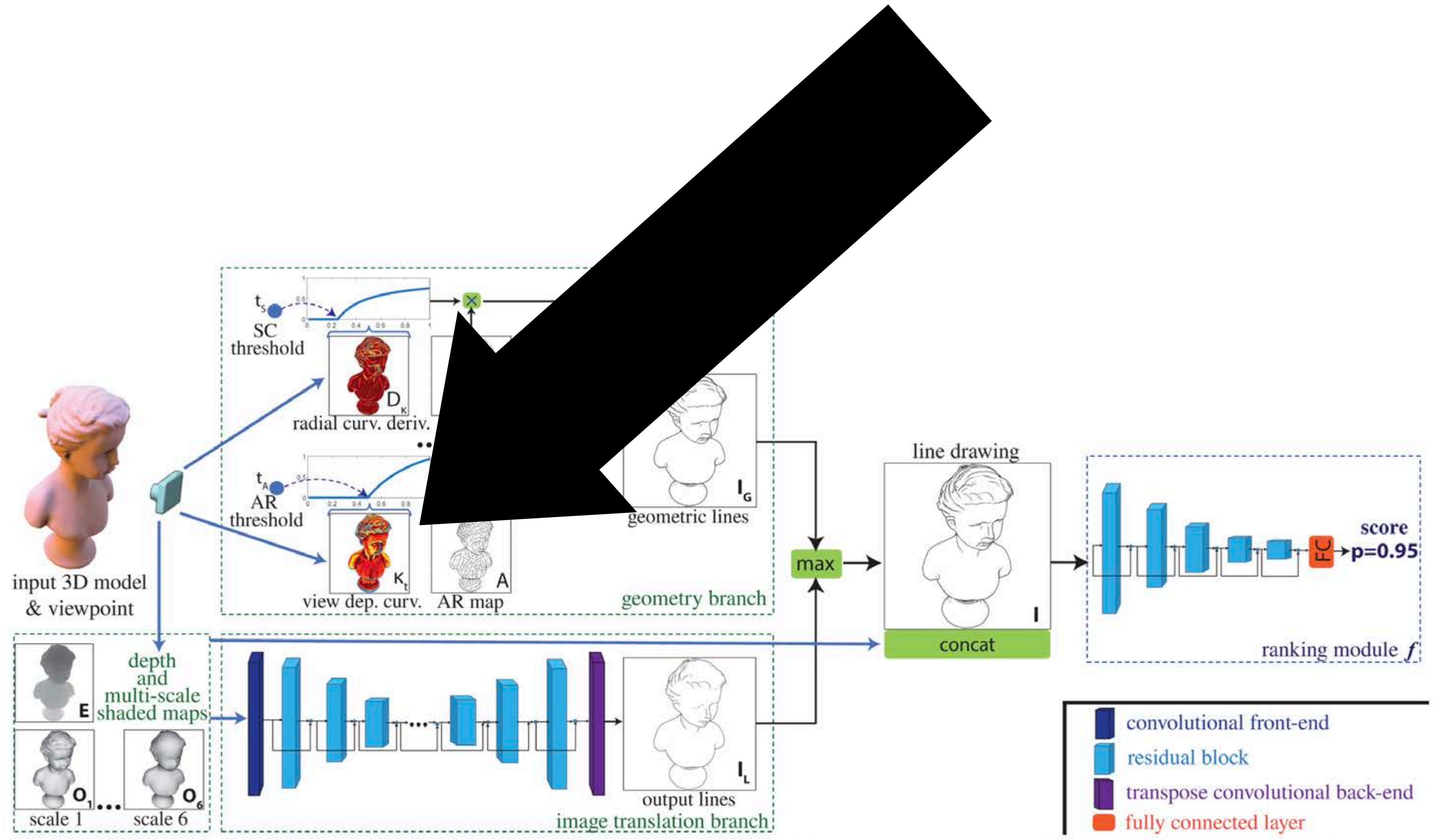


Figure 3: Our network architecture: the input 3D model is processed by a geometry branch operating on curvature features, and an image-based branch operating on view-based representations. Their outputs are combined to create a line drawing, which is in turn evaluated by a ranking module that helps determining optimal line drawing parameters.

Apparent Ridges as inverse of inverse?

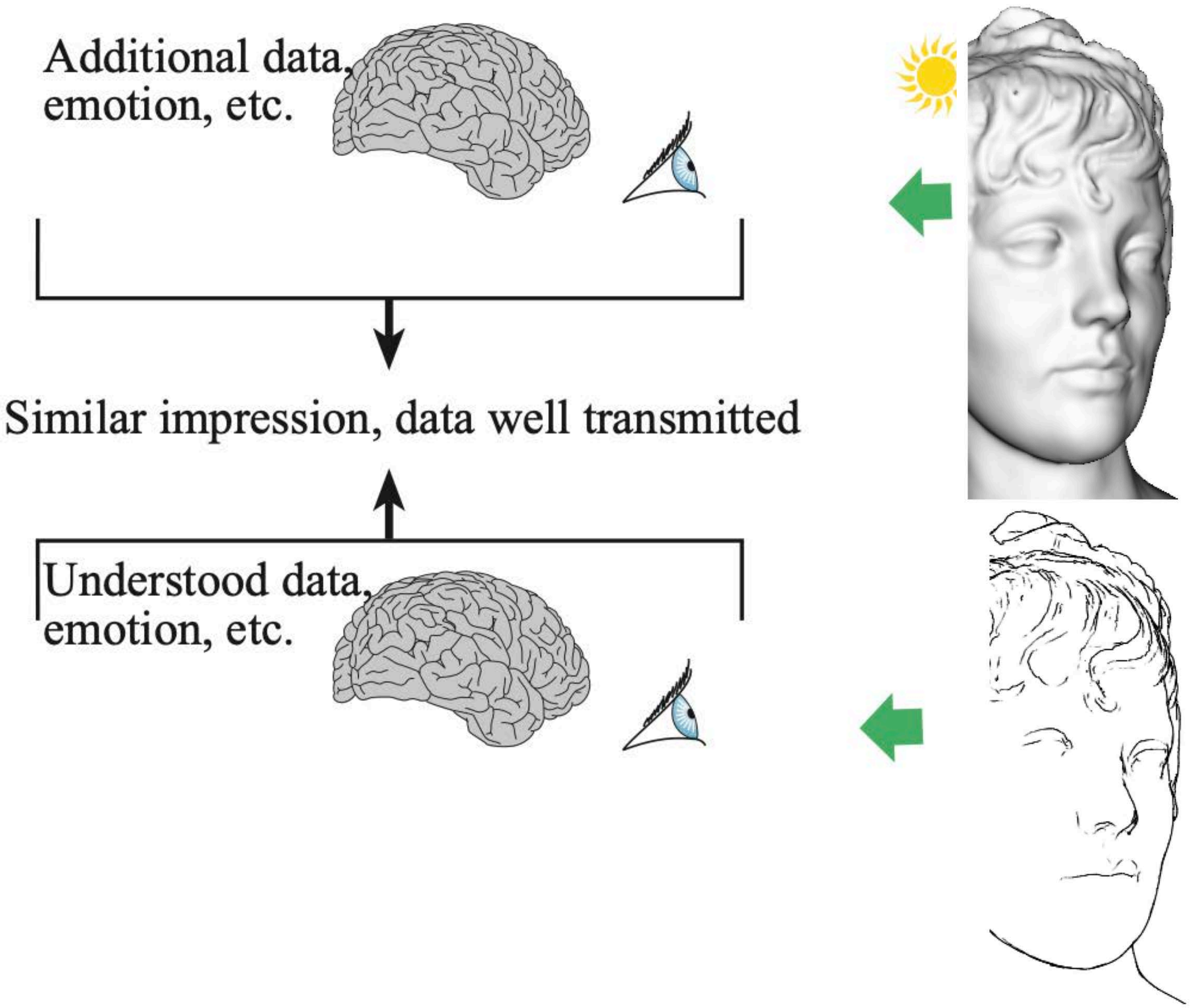


Figure 6: Depiction as the inverse of an inverse problem.

Apparent Ridges as inverse of inverse?

Requires Ted's brain to provide insights

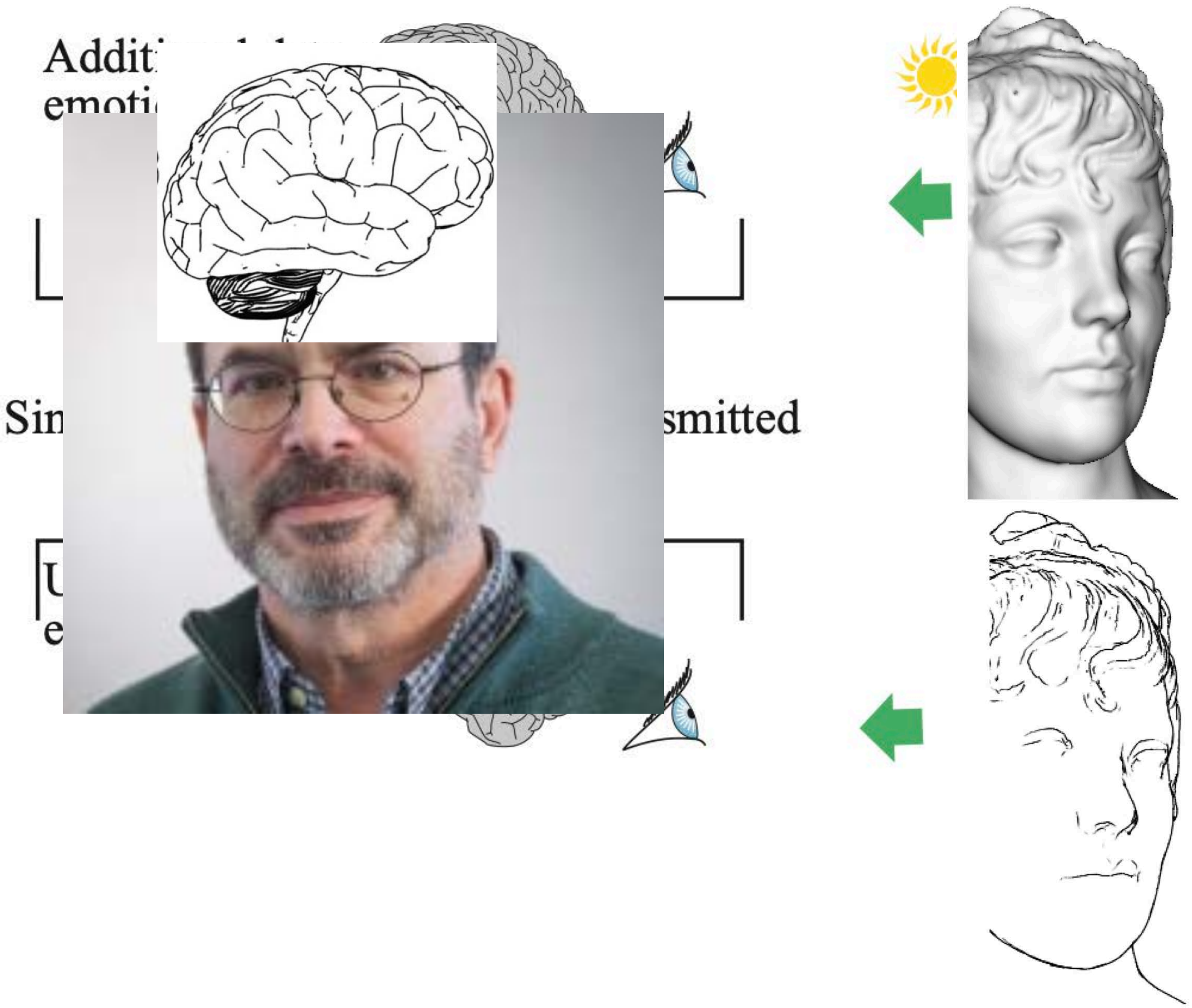
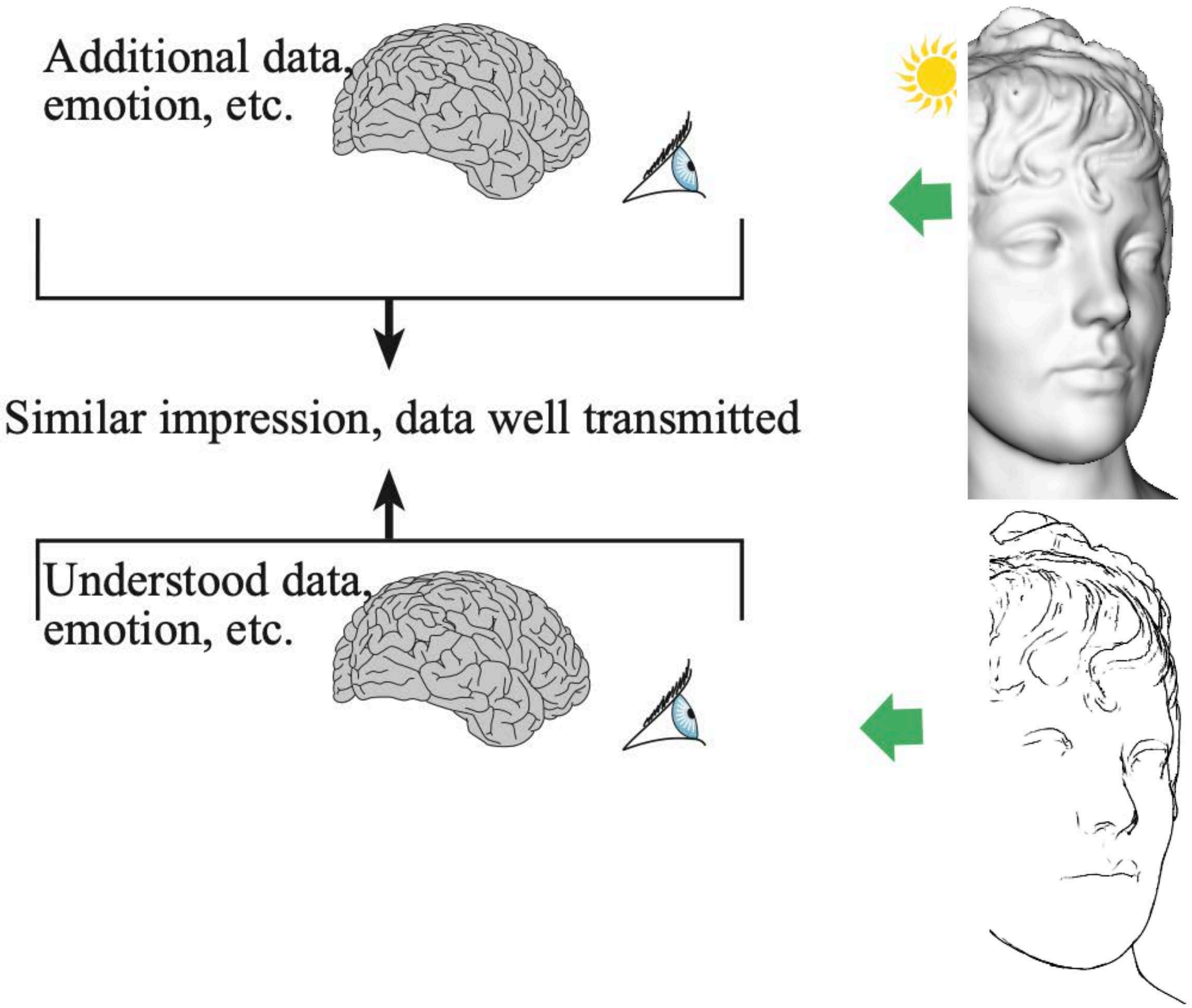


Figure 6: Depiction as the inverse of an inverse problem.

Can we generate line drawings by optimizing the similarity of the percept they elicit to the real percept?

Put human brains in an optimization inner loop?



What if we replace the brain by an artificial brain?

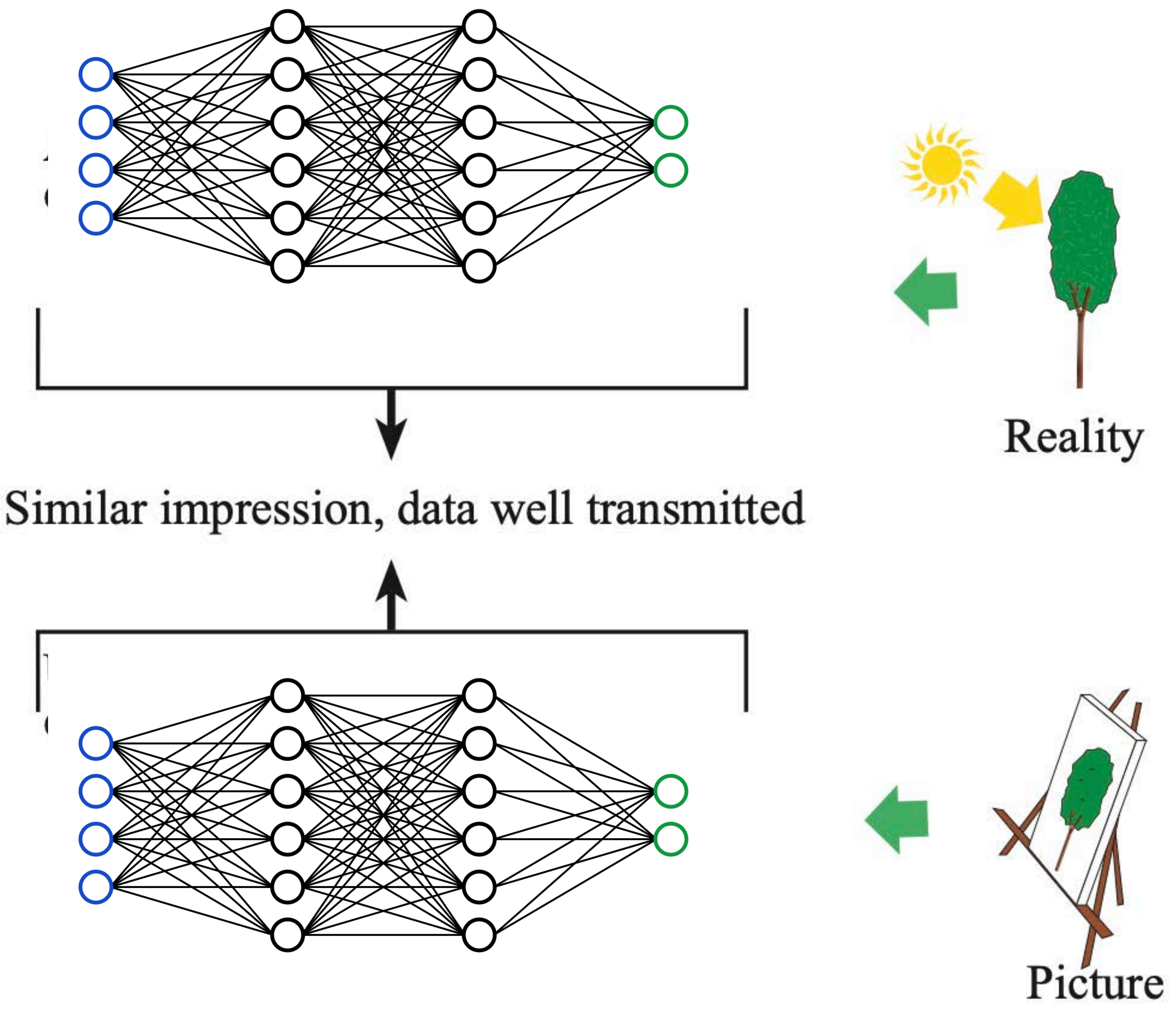


Figure 6: Depiction as the inverse of an inverse problem.

Informative line drawings

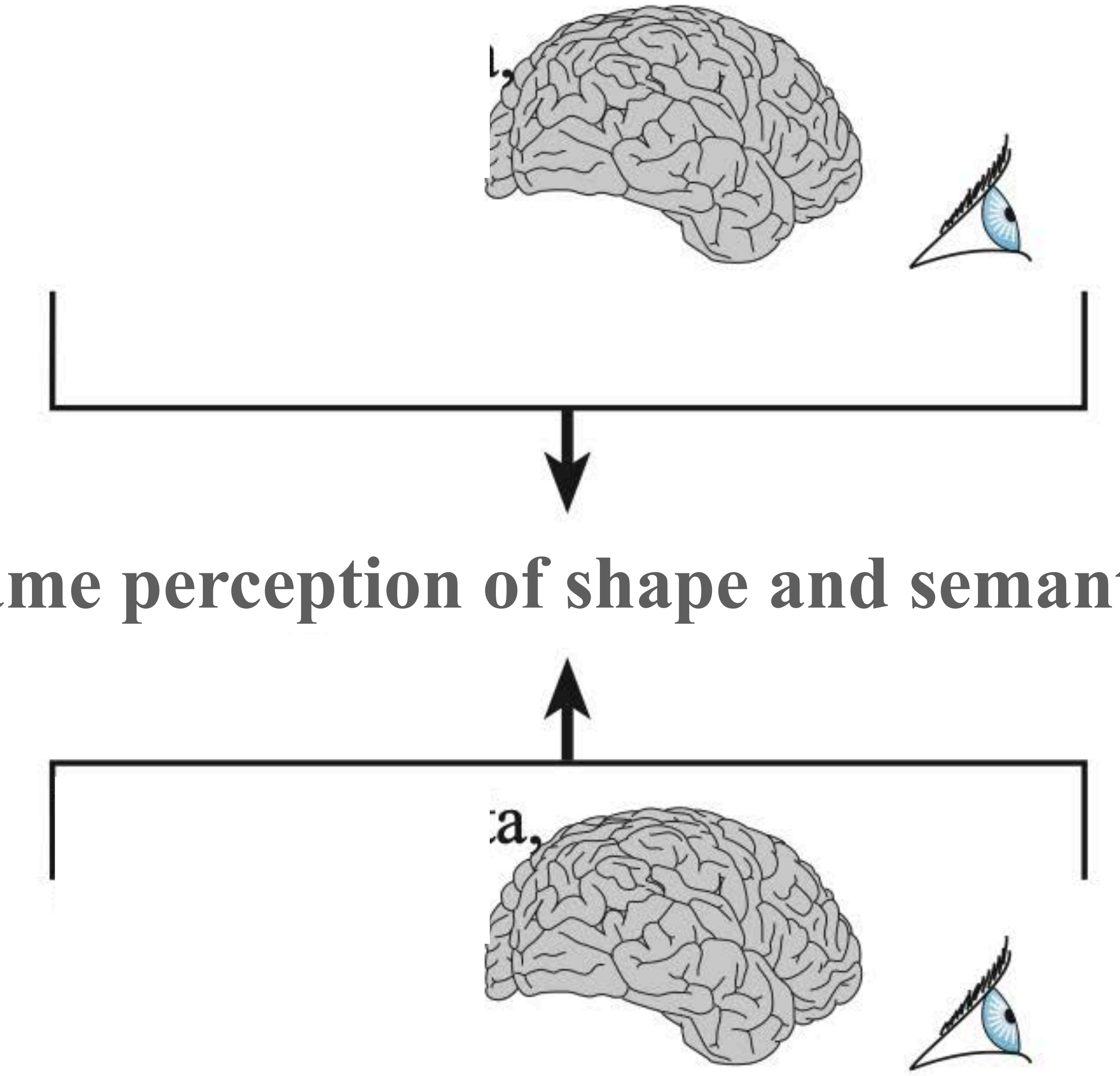
Learning to generate line drawings that convey geometry and semantics

Caroline Chan Frédo Durand Phillip Isola
{cmchan, fredo, phillipi}@mit.edu
MIT

Figure 1. Given a set of photographs, our method is capable of making line drawings in different styles seen above. Our method only requires unpaired data during training.

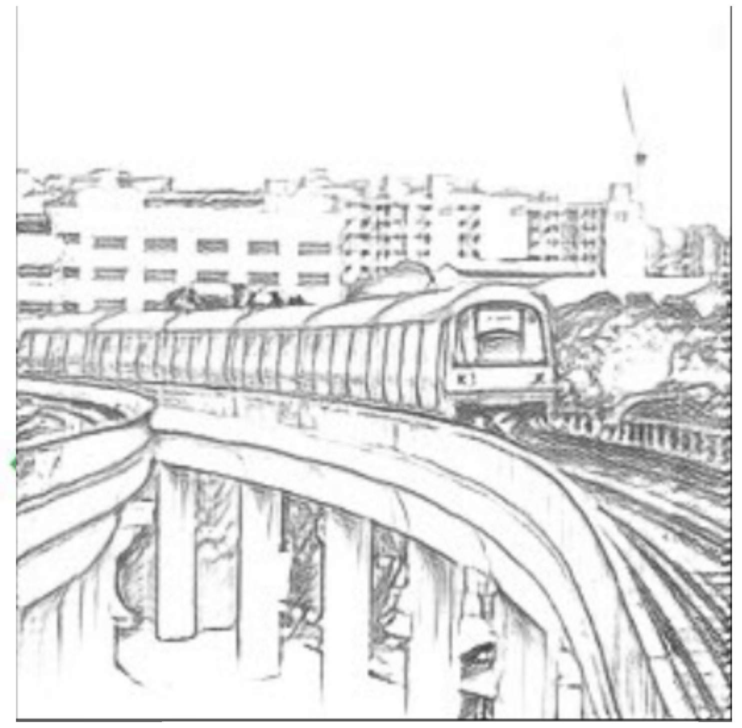
Goal: generate line drawing that can convey the same 3D shape and somatic as an input photo

Without paired training data



Same perception of shape and semantics

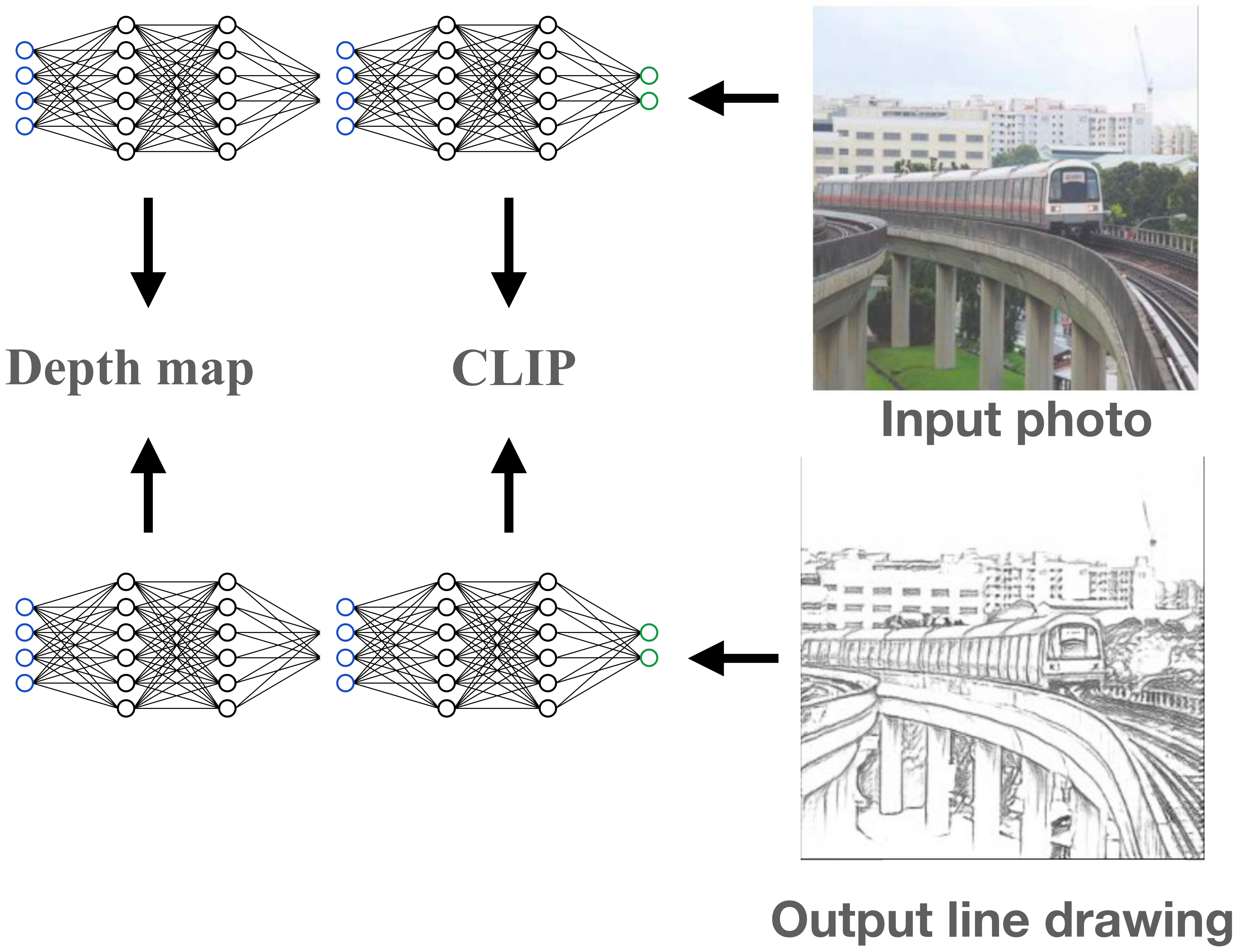
Input photo



Output line drawing

Approach: compare prediction for shape (depth) and semantics (CLIP) for photo & output drawing

Plus compare RGB, and a GAN loss on unpaired training drawings



+ Style Loss

Approach: compare prediction for shape (depth) and semantics (CLIP) for photo & output drawing

Plus compare RGB, and a GAN loss on unpaired training drawings

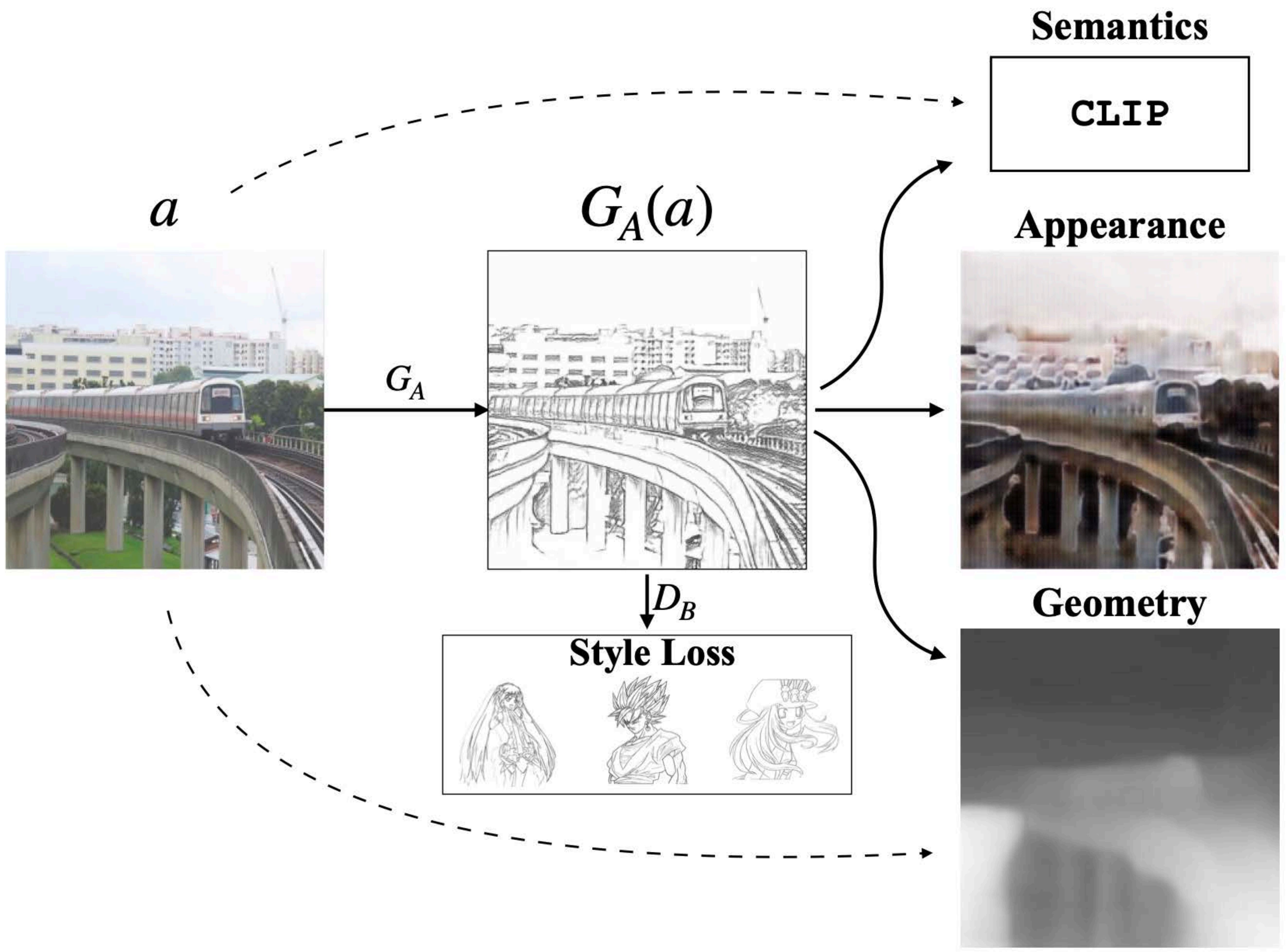
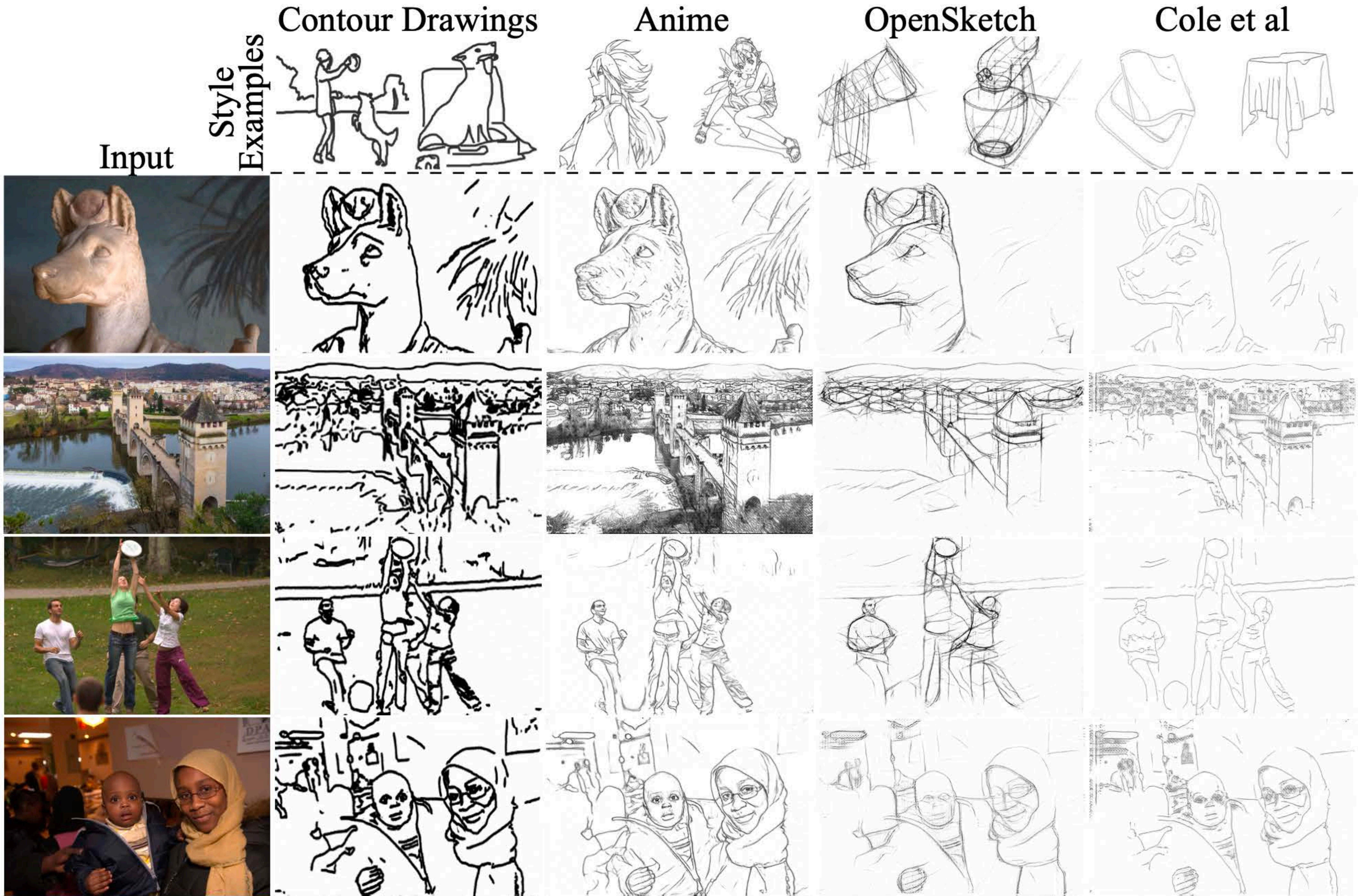


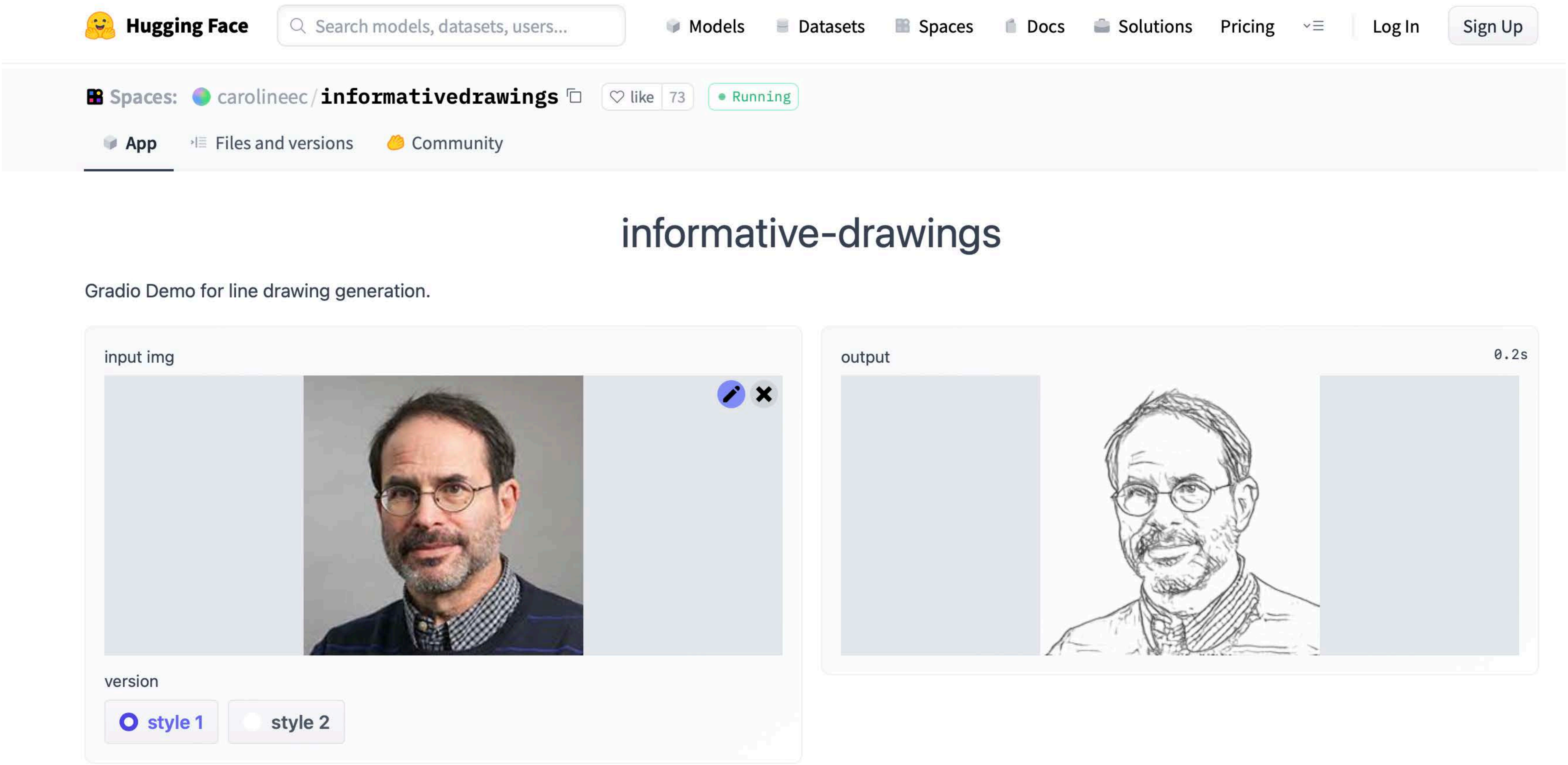
Figure 2. Given a photograph a, our model trains network G_A to synthesize line drawing $G_A(a)$ via four main losses. Adversarial style loss with discriminator D_B encourages generated line drawings to match the style of the training set. The CLIP, appearance, and geometry losses enforce that the line drawing communicates effective semantic, appearance, and geometry respectively.


Input
CycleGAN
TSIT
U-GAT-IT
SPatchGAN
Council-GAN
ACL-GAN
UPDG
Ours


Contour Drawings
Anime
OpenSketch
Cole et al
Style Examples
Input

Try it!

https://huggingface.co/spaces/carolineec/informativedrawings



Thank you Ted

Some of the things I have learned

- Look for questions. Explore half-formed questions
- Explore, experiment, in your head or with small models

- Step back, is the big picture making sense
- Don’t be afraid to work with people with more expertise than you in various areas
- Multidisciplinary work keeps life interesting
- Art matters

- Trust your own visual system.
- You don’t need a big research group to have impact
- One piece of great work is more important than multiple good works (quality not quantity)
- Don’t let sponsors make your life difficult