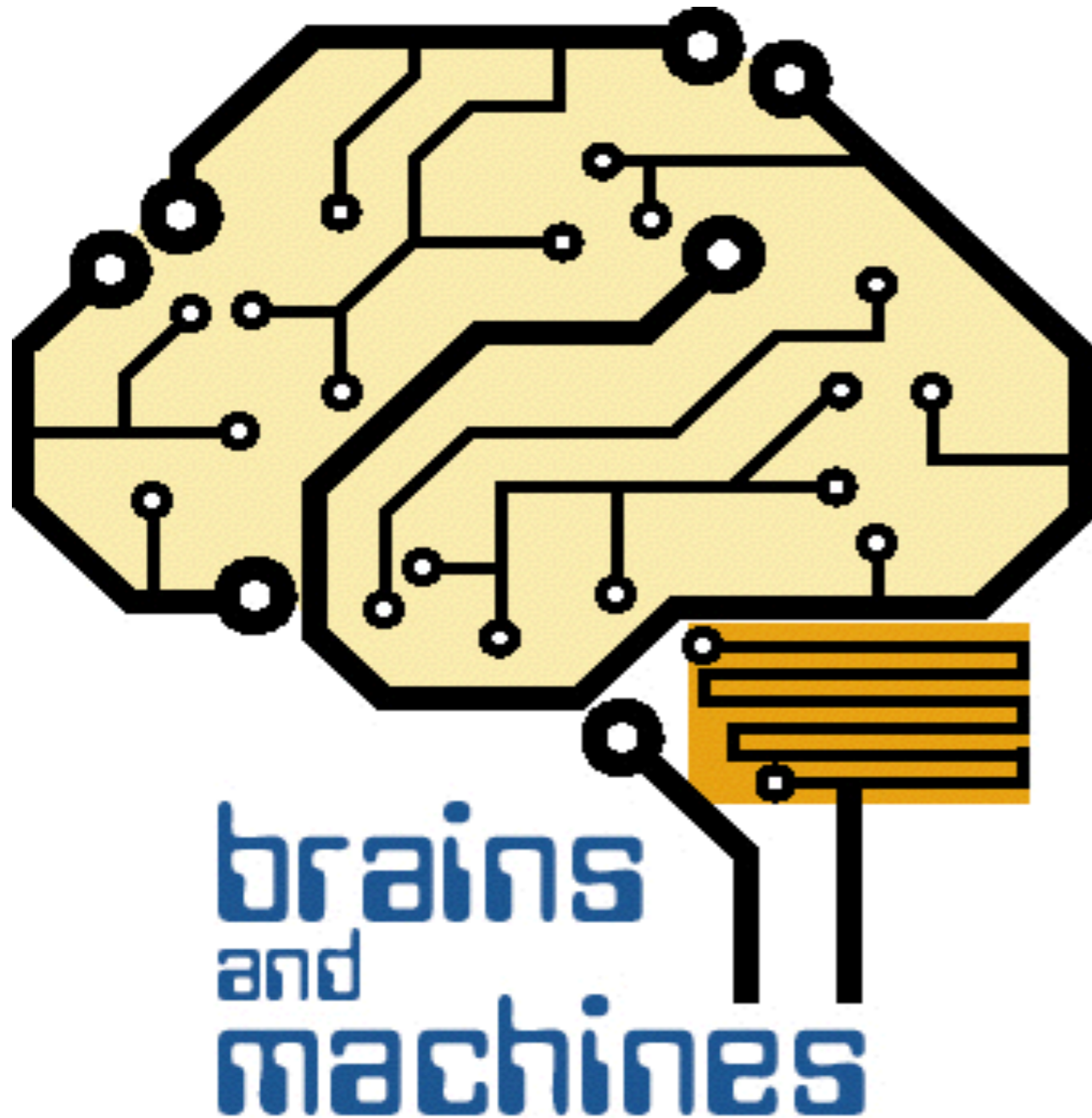
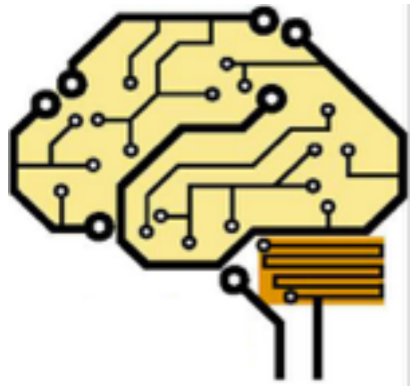


How visual cortex works and machine vision tools for the visually impaired



tomaso poggio
CBMM,
McGovern Institute, BCS, CSAIL
MIT

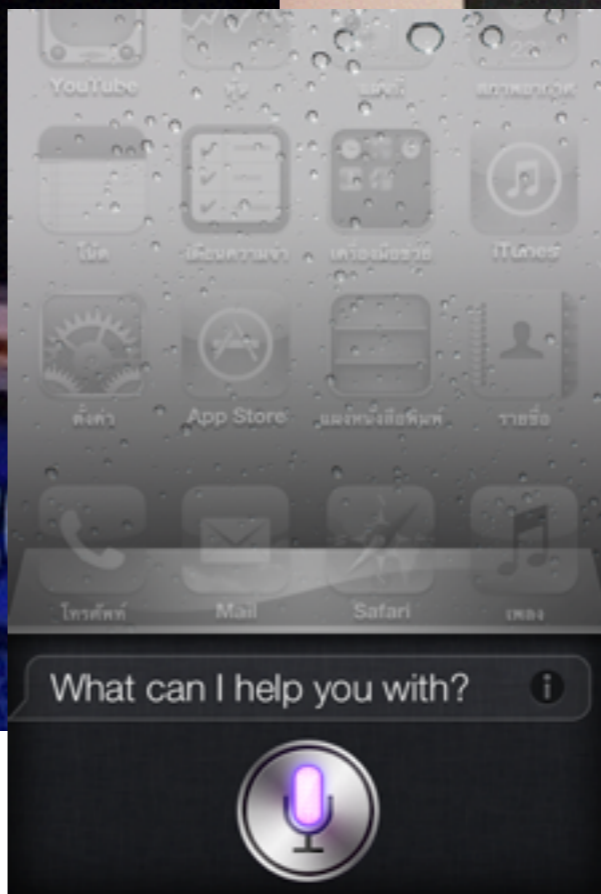
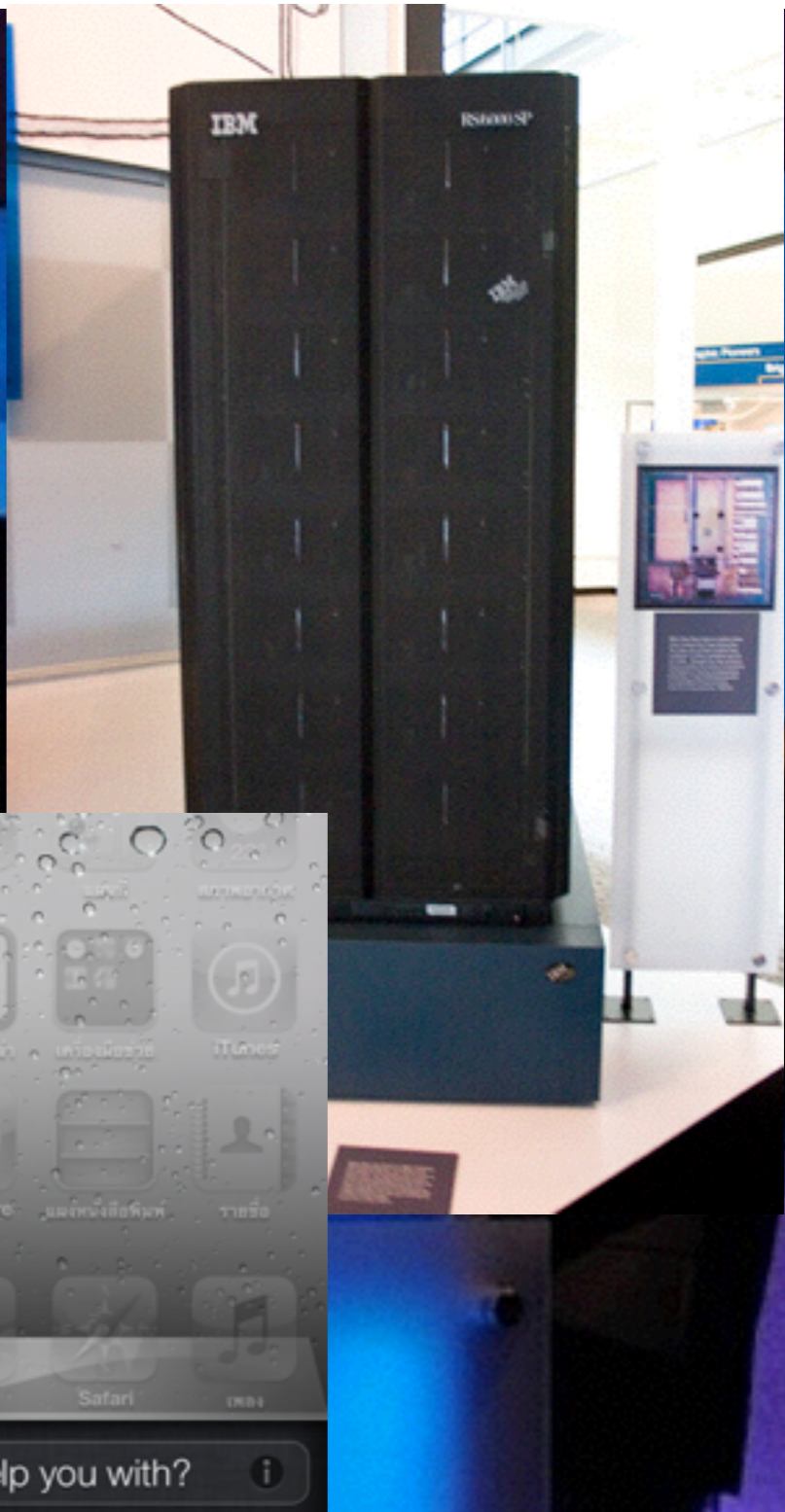
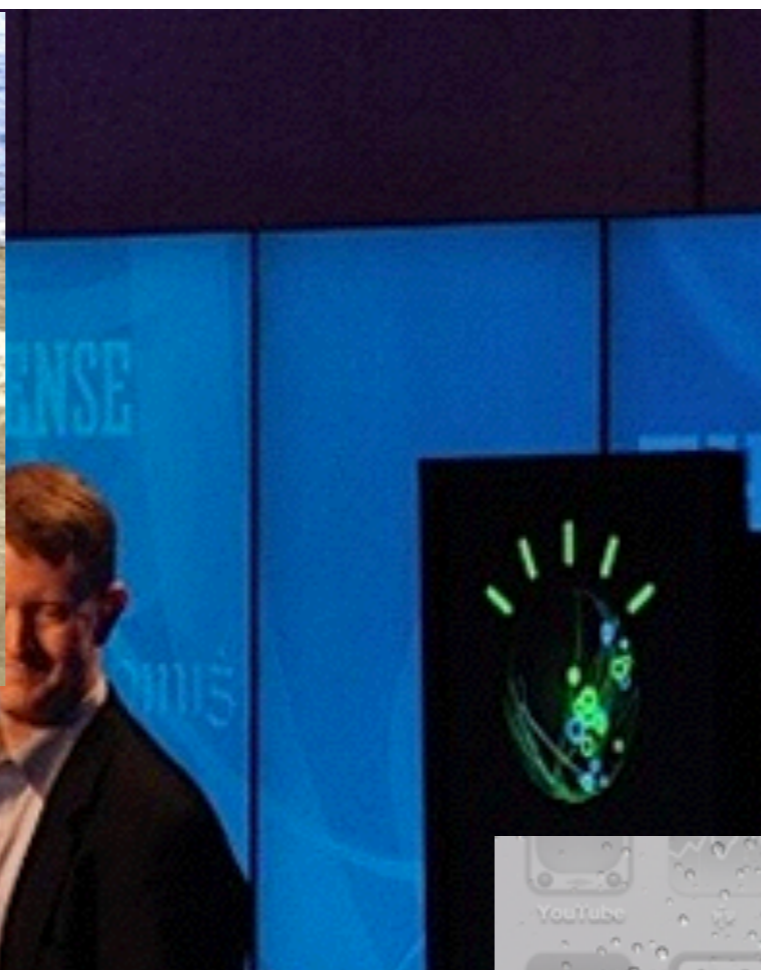


The Center for Brains, Minds and Machines



Vision for CBMM

- The problem of intelligence is one of the great problems in science.
- Work so far has led to many systems with impressive but narrow intelligence
- Now it is time to develop a basic science understanding of human intelligence so that we can take intelligent applications to another level.



Pedestrian accidents occur every day
in our increasingly intensive traffic environment.
In Europe, 14% of all traffic fatalities are pedestrians.

MIT

Boyden, Desimone ,Kaelbling , Kanwisher,
Katz, Poggio, Sassanfar, Saxe,
Schulz, Tenenbaum, Ullman, Wilson,
Rosasco, Winston

Harvard

Blum, Kreiman, Mahadevan,
Nakayama, Sompolinsky,
Spelke, Valiant

Cornell

Hirsh

Allen Institute

Koch

Rockefeller

Freiwald

UCLA

Yuille

Stanford

Goodman

Hunter

Epstein,...

Wellesley

Hildreth, Conway...

Puerto Rico

Bykhovaskaia, Vega...

Howard

Manaye,...



City U. HK

Smale

Hebrew U.

Shashua

IIT

Metta, Rosasco,
Sandini

MPI

Buelthoff

NCBS

Raghavan

Genoa U.

Verri

Weizmann

Ullman



Google

Norvig

IBM

Ferrucci

Microsoft

Blake

Orcam

Shashua

MobilEye

Shashua

DeepMind

Hassabis

**Boston
Dynamics**

Raibert

**Rethink
Robotics**

Brooks

**Willow
Garage**

Cousins



BostonDynamics



Rational for a Center

Convergence of progress: a key opportunity

Machine Learning &
Computer Science

Neuroscience &
Computational
Neuroscience

Cognitive Science

*Change comes most of all from the unvisited no-man's-land
between the disciplines (Norbert Wiener (1894-1964))*

Science + Technology
of Intelligence

The core CBMM challenge: describe in a scene
objects, people, actions, social interactions

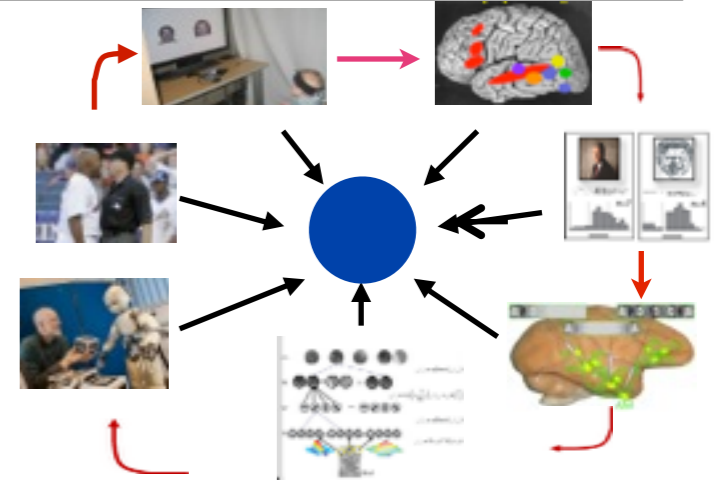
What is this?

What is Hueihan doing?

What does Hueihan think about Joel's thoughts about her?



The core CBMM challenge: measuring progress



The tests include an open set of questions and answers on databases of images and videos about physical and social world:

- *what is there*
- *who is there*
- *what is the person doing*
- *who is doing what to whom?*
- *what happens next?*

We will measure our progress over the life of the CBMM by evaluating a) how well our neural models of the brain answer the questions, b) how similarly to humans c) how much their implied physiology correlates with human and primate data in these tasks.

City U. HK

Smale

Hebrew U.

Shashua

IIT

Metta, Rosasco,
Sandini

MPI

Buelthoff

NCBS

Raghavan

Genoa U.

Verri

Weizmann

Ullman



Google

Norvig

IBM

Ferrucci

Microsoft

Blake

Orcam

Shashua

MobilEye

Shashua

DeepMind

Hassabis

**Boston
Dynamics**

Raibert

**Rethink
Robotics**

Brooks

**Willow
Garage**

Cousins



OrCam

Harnessing Computer Science to aide the
Visually Impaired

Visual Acuity

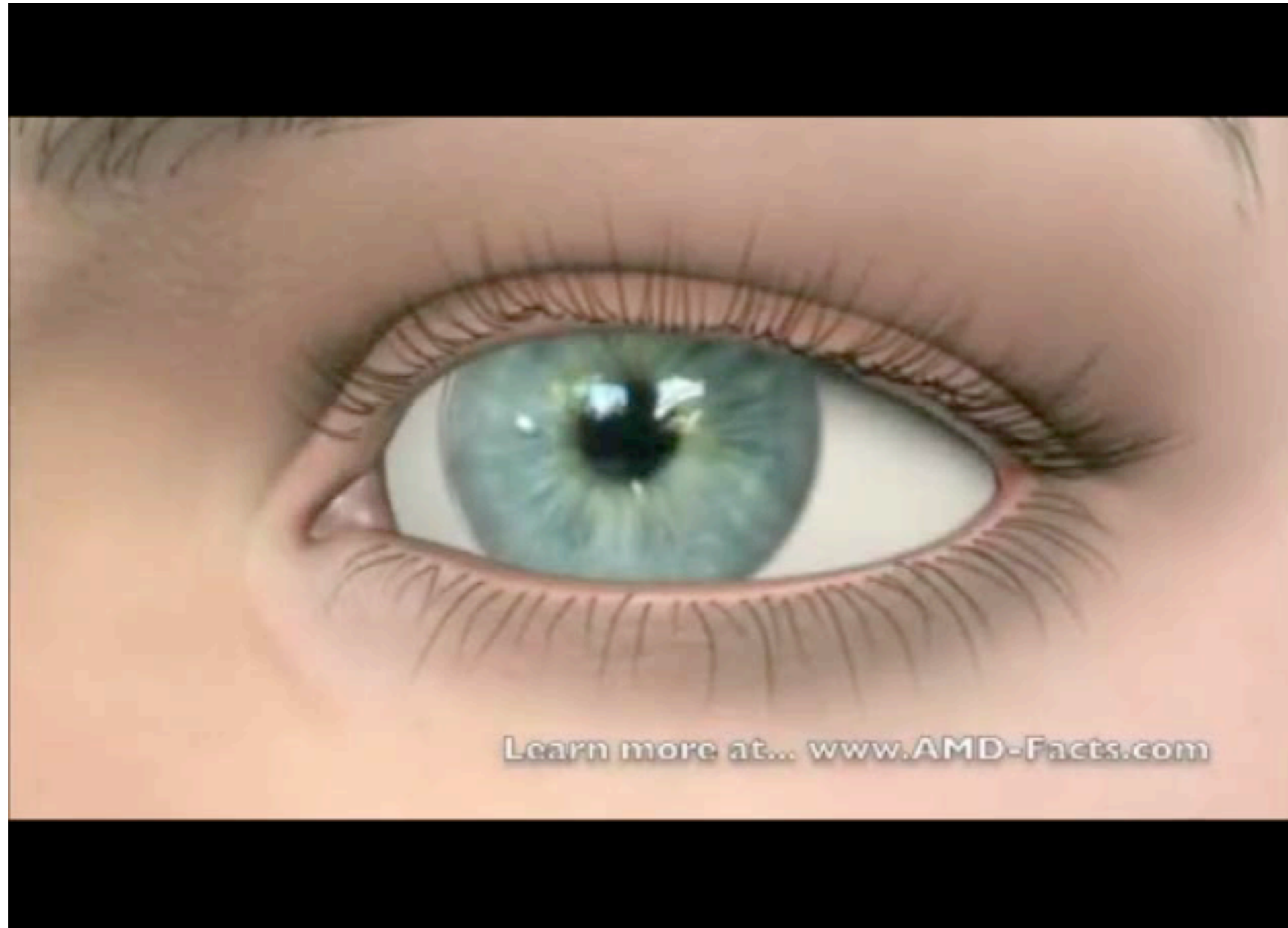
Legally blind	$\frac{20}{200}$	E	$\frac{200 \text{ FT}}{61 \text{ M}}$	1
Low Vision (with glasses)	$\frac{20}{100}$	F P	$\frac{100 \text{ FT}}{30.5 \text{ M}}$	2
	$\frac{20}{70}$	T O Z	$\frac{70 \text{ FT}}{21.3 \text{ M}}$	3
Visually Impaired (with glasses)	$\frac{20}{50}$	L P E D	$\frac{50 \text{ FT}}{15.2 \text{ M}}$	4
	$\frac{20}{40}$	P E C F D	$\frac{40 \text{ FT}}{12.2 \text{ M}}$	5
	$\frac{20}{30}$	<u>E D F C Z P</u>	$\frac{30 \text{ FT}}{9.14 \text{ M}}$	6
No need for glasses	$\frac{20}{25}$	F E L O P Z D	$\frac{25 \text{ FT}}{7.62 \text{ M}}$	7
	$\frac{20}{20}$	<u>D E F P O T E C</u>	$\frac{20 \text{ FT}}{6.10 \text{ M}}$	8
	$\frac{20}{15}$	L E F O D P C T	$\frac{15 \text{ FT}}{4.57 \text{ M}}$	9
	$\frac{20}{13}$	F D P L T C E O	$\frac{13 \text{ FT}}{3.96 \text{ M}}$	10
	$\frac{20}{10}$	F R E S O L C E T O	$\frac{10 \text{ FT}}{3.05 \text{ M}}$	11

In the expression, **20/X vision**, the 20 is the distance in feet between the subject and the chart. The X means that the subject can read the chart (from 20 feet away) as well as a normal person could read the same chart from X feet away.

Visual Impairment

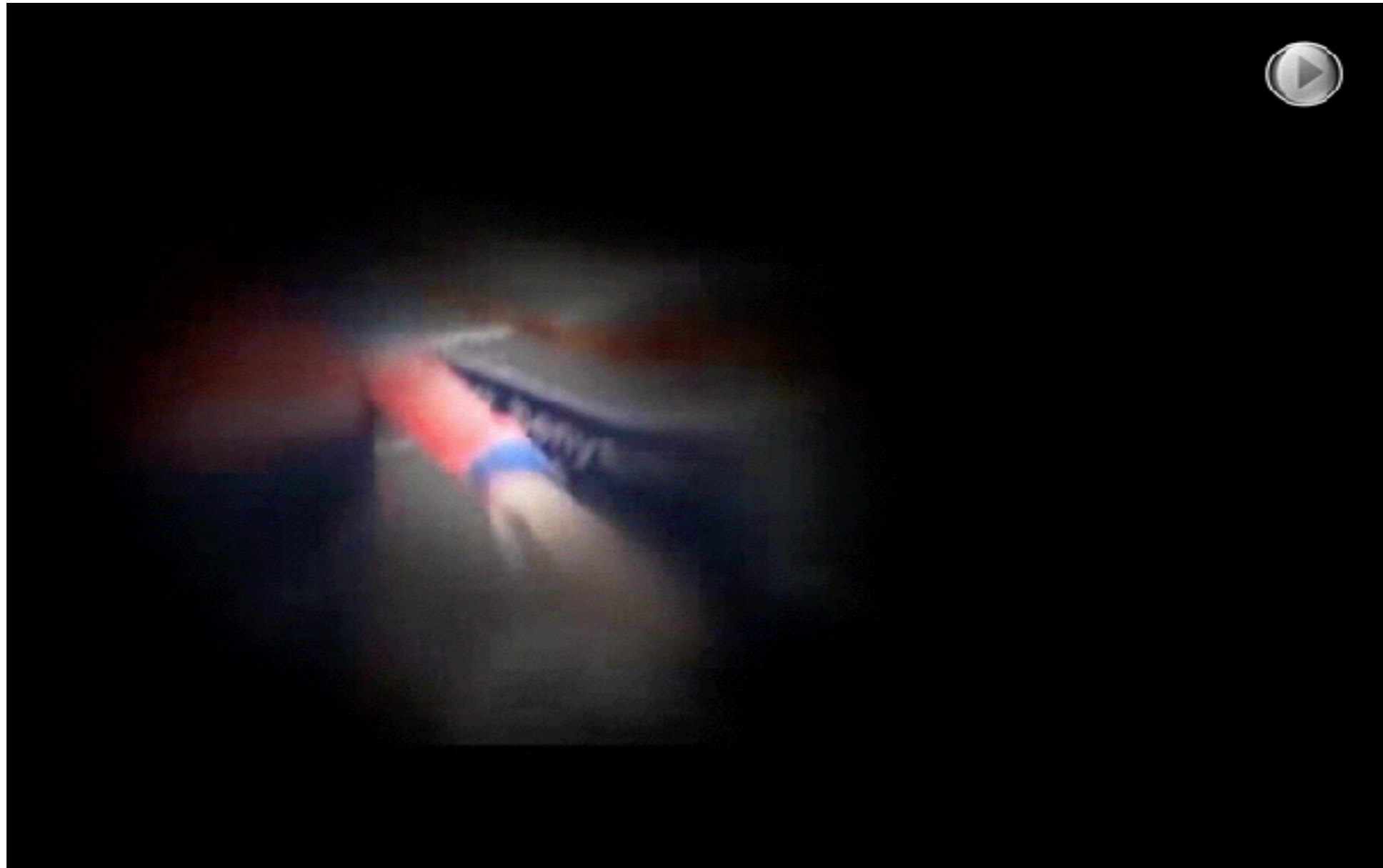
- Age-related Macular Degeneration
- Retinitis Pigmentosa
- Diabetic retinopathy
- Macular Edema-Ischemia
- Coloboma (birth defect, 1 / 10,000 people)
- .
- .

How does it feel?



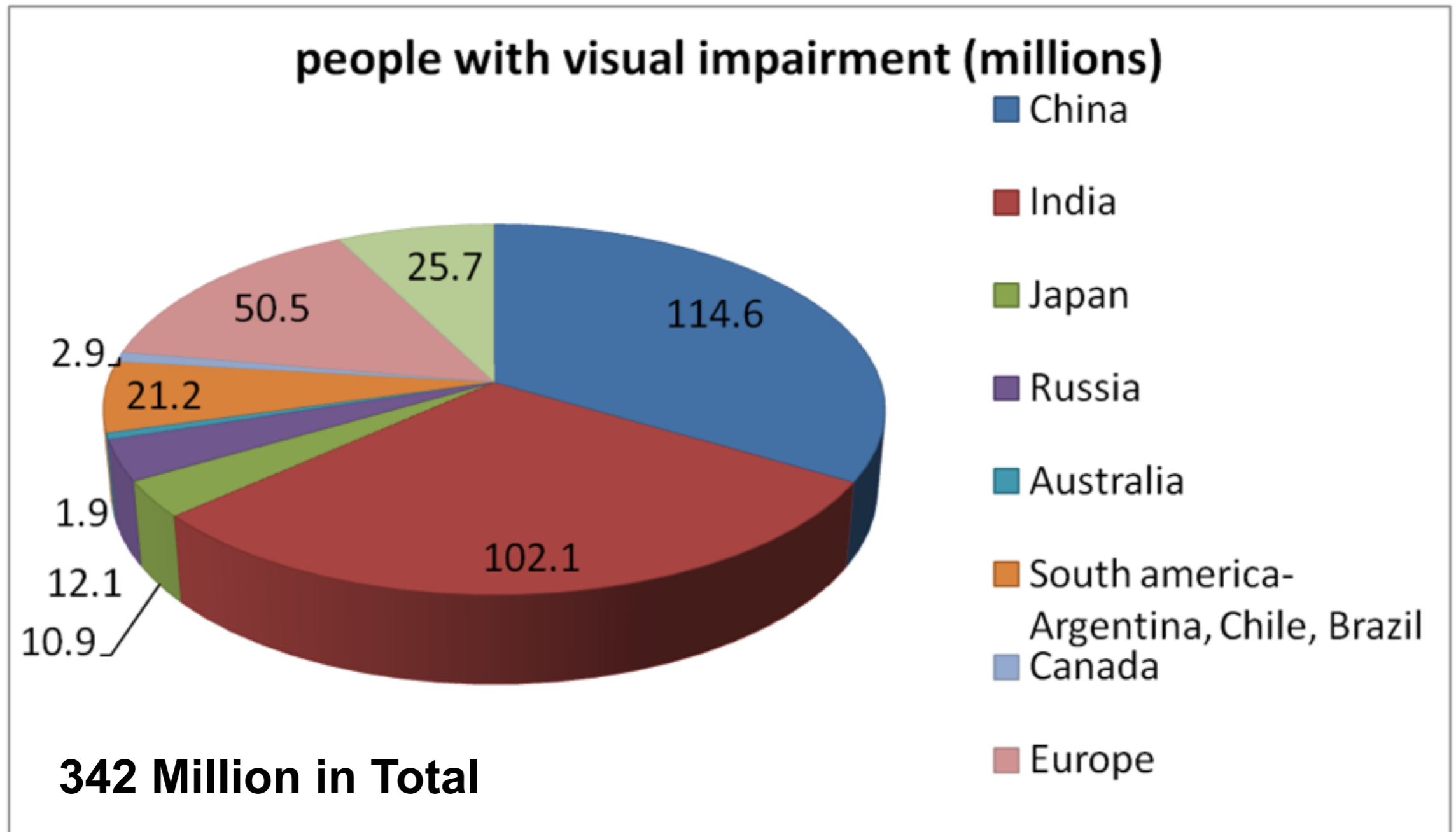
AMD - Age Macular Degeneration
ניוון מקולרי גילאי

How does it feel?



Retinitis Pigmentosa

Prevalence in the World



•Based on CIA world fact book data

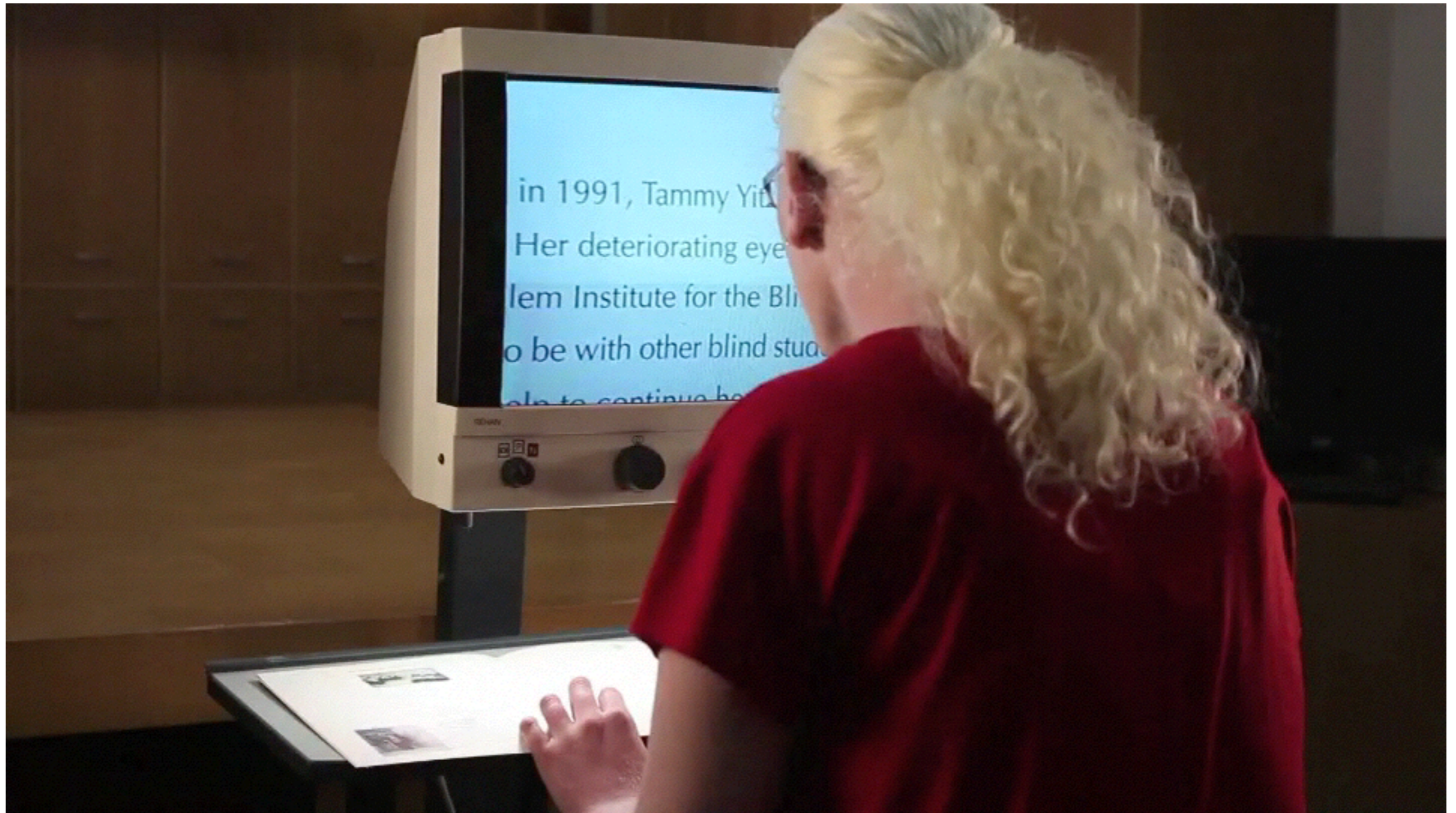
How Does it

Simple Daily Activities are a Challenge



The visual world is rich with textual information: billboards, street names, newspapers, books,.. Then there is getting around: crossing a street, catching a bus, choosing a product, paying for it...

Assistive Technologies

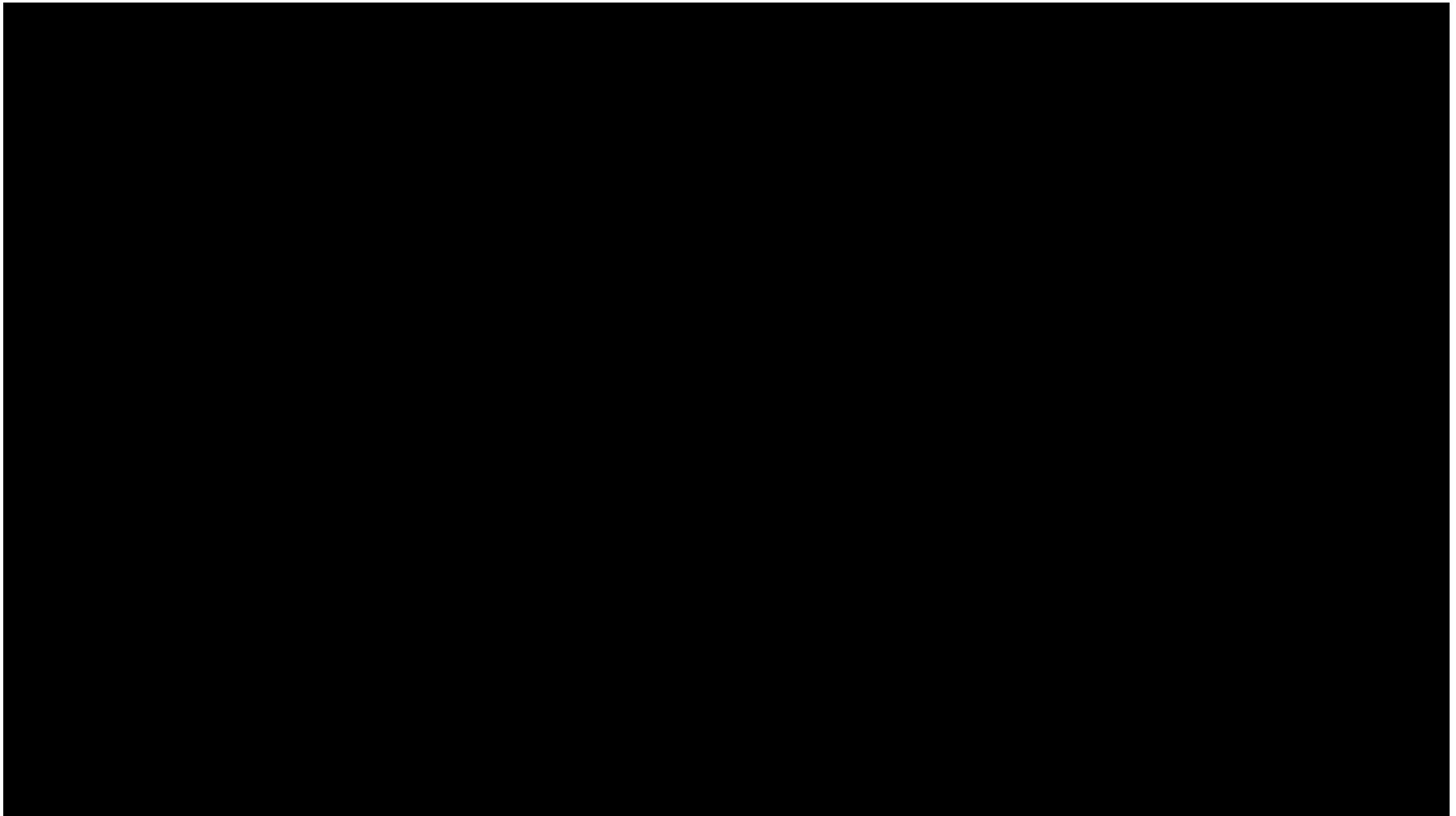


Assistive Technologies

PRICE (App.)	FUNCTION	DEVICE
Free	OCR of documents Read Money Notes (limited) object recognition (via cloud computing)	Smartphone apps
800\$	OCR of documents	Intel Re



Assistive Technologies



“SayText: an iPhone App for the visually Impaired

Assistive Technologies



LookTel OCR

What Does



OrCam Product

Learns!



Reads



Recognizes Faces



Recognizes Places



Recognizes Objects





Reading



Document

Lorem ipsum dolor sit amet.
Consectetur adipiscing elit. Cras non nunc nec
enim tristique tincidunt. Vestibulum quis felis.
Duis nulla. Donec lectus urna. Sed tempus nibh
id massa. Vivamus placerat justo quis nibh. Ut
quis ante. Ut scelerisque quam eu mi. Donec
morisae purus sit amet velit. Sed ac sem.
Aenean quis justo. Vestibulum ante ipsum primis
in faucibus orci luctus et ultrices posuere cubilia
Curae; Ut tincidunt.
Nulla facilis. Aenean eros felis, blandi eu,
conmodo sit amet, varius a, pede. Curabitur
augue felis, congue sed.

Recognizes Objects



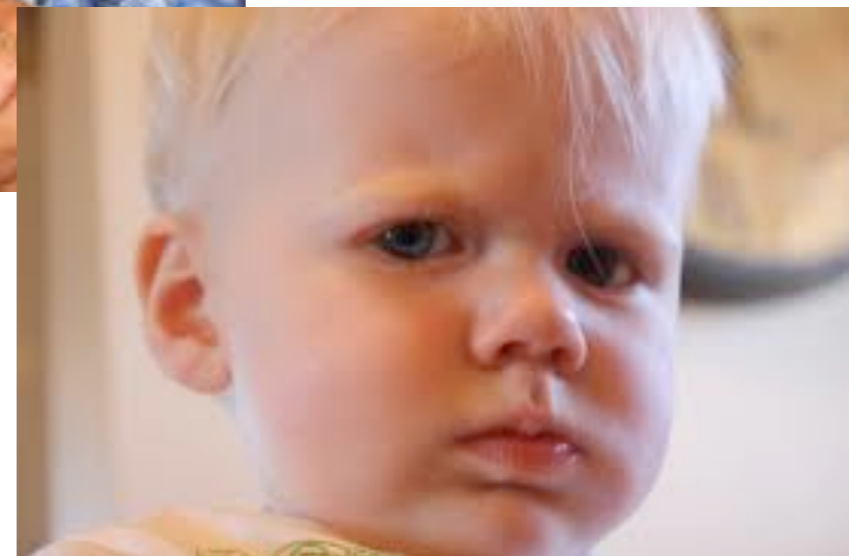
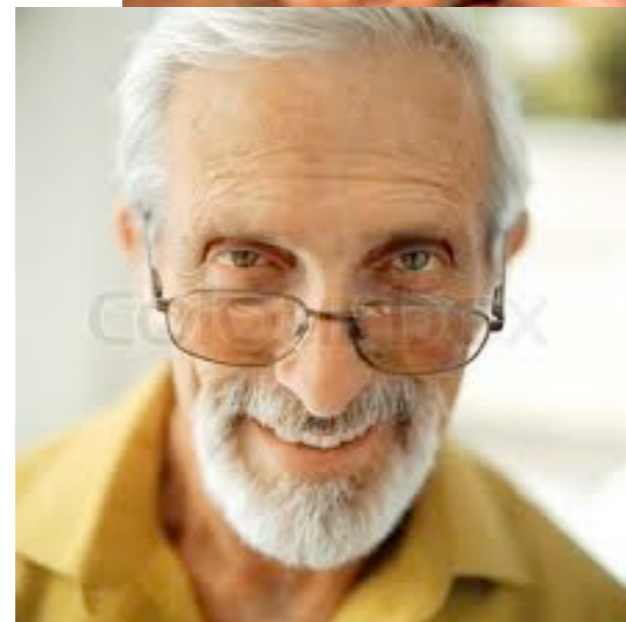
Learns New Objects

Preloaded hundreds of basic products

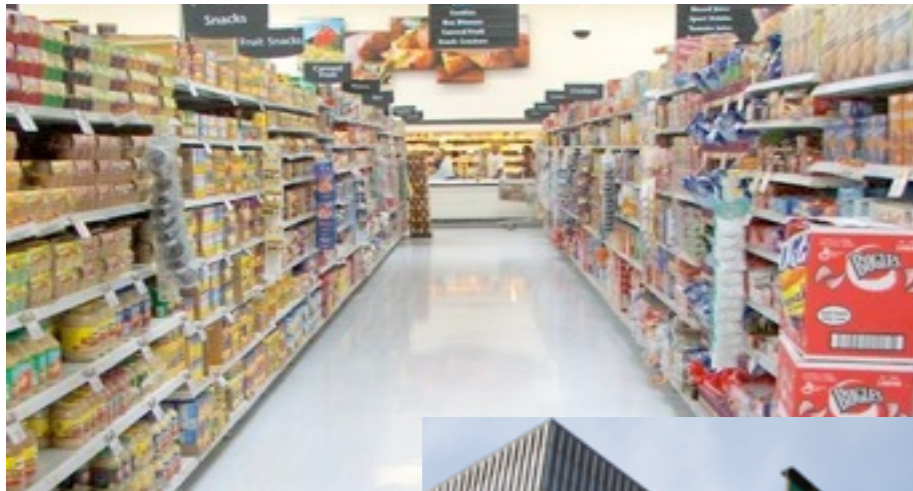




Learns and Recognizes



Learns and Recognizes

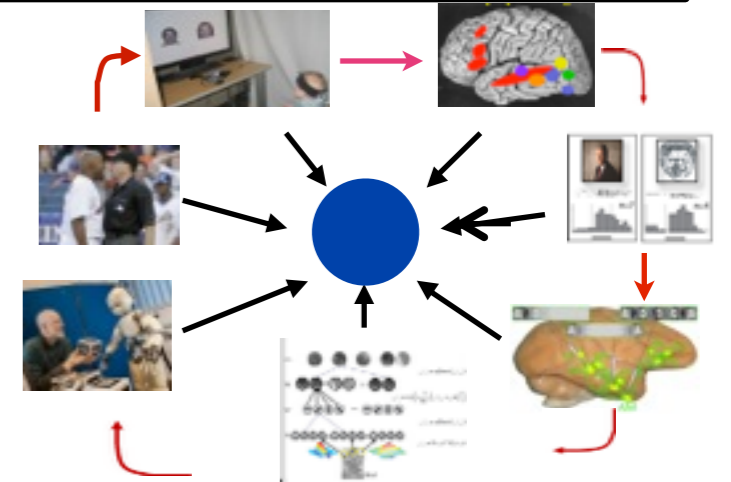


User Trials





The core CBMM challenge: human-like visual understanding



The tests include an open set of questions and answers on databases of images and videos about physical and social world:

- *what is there*
- *who is there*
- *what is the person doing*
- *who is doing what to whom?*
- *what happens next?*

We will measure our progress over the life of the CBMM by evaluating a) how well our neural models of the brain answer the questions, b) how similarly to humans c) how much their implied physiology correlates with human and primate data in these tasks.