

# ESP游戏

徐迎晓

[xuyingxiao@126.com](mailto:xuyingxiao@126.com)



# ESP Game

- A new interactive system in the form of a game with a unique property: the people who play the game label images for us.
- 用途
  1. 有视觉障碍的人上网时可听到对图片的描述
  2. 研究计算机视觉者可用大量有标签的图片训练机器学习算法
  3. 图片搜索
  4. 图片过滤
  5. 个性化推荐 (New)
- 效果:
  - 5,000 people continuously playing the game could assign a label to all images indexed by Google in 31 days.



Player 1 guesses: purse  
Player 1 guesses: bag  
Player 1 guesses: brown

Success! Agreement on “purse”



Player 2 guesses: handbag

Player 2 guesses: purse  
Success! Agreement on “purse”

ESP (Extra Sensual Perception)

**0:11**  
Time Left

# The ESP Game

**2100**  
score



## Taboo Words

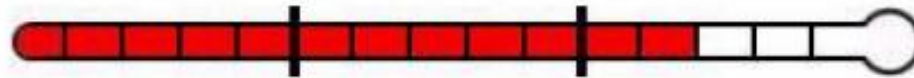
MAN  
BEARD

## Your Guesses

HAT

Type your next guess:

Pass



© 2002-2003 Carnegie Mellon University, all rights reserved. Patent Pending.

# Design Issues

- Incremental success
- Pass
- Diversity
- Ensure the quality of the labels

**When is an Image “Done”?**

玩家数目不够？

- **Cheating**, 系统级解决方法是:
  1. 将不同地方的人配在一起, 随机匹配, 碰巧匹配到身边的人概率小
  2. 记录IP地址, 匹配者的IP地址必须不在不同网段
  3. 系统检测到存在massive agreement strategy后, 放入大量的pre-recorded game-play, 等尝试cheating者放弃后, 再逐步减少pre-recorded game-play
    1. 如何检查massive agreement strategy? —— 计算平均时间



# Selecting the Images

- blank images
- images that consist of a single color,
- images that are smaller than 20 pixels on either dimension,
- images with an aspect ratio greater than 4.5 or smaller than  $1 / 4.5$

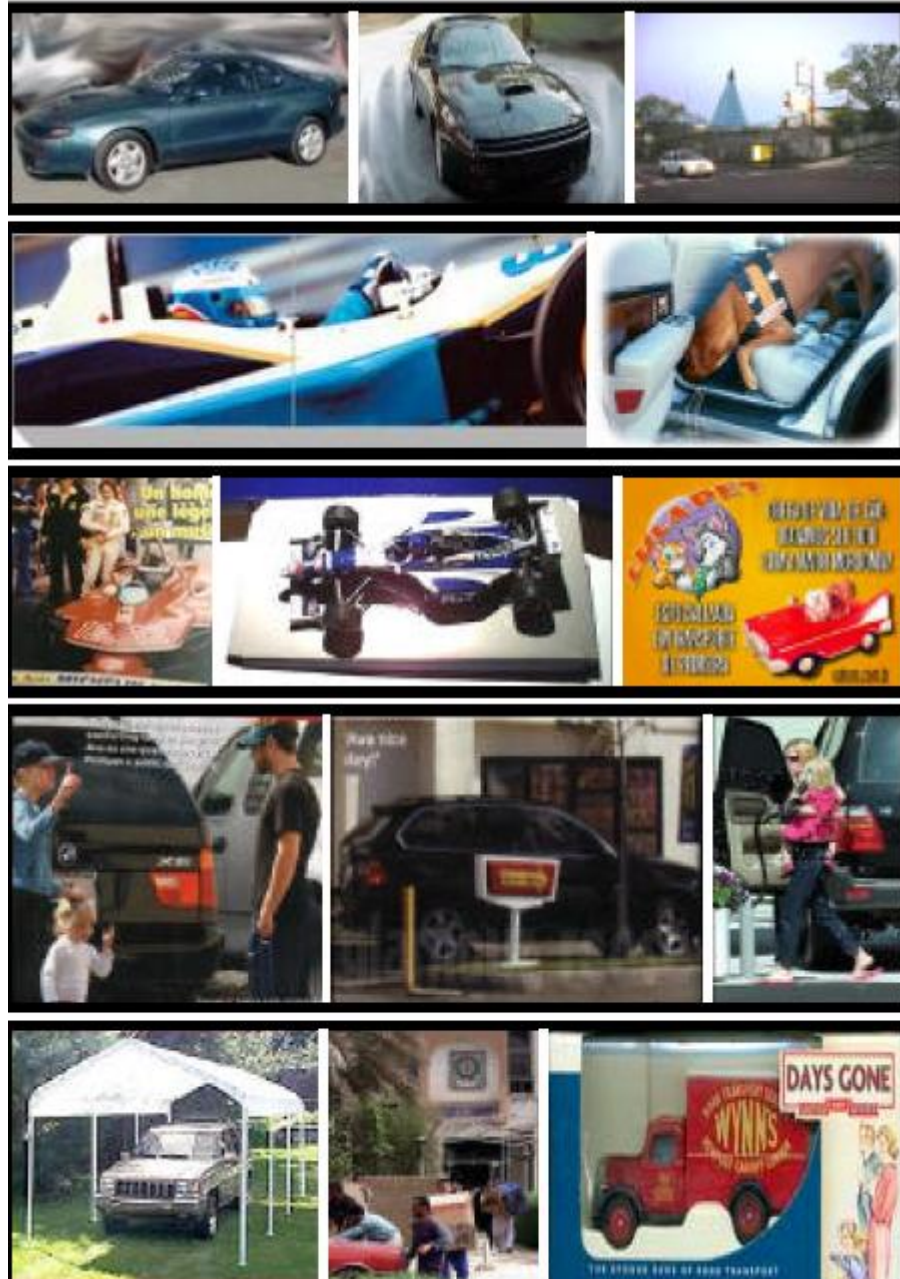
- Misspelling
- **Context-Specific Labels**
- **Inappropriate Content**

# EVALUATION

- 如何证明玩家喜欢？
- 如何证明Label的质量好？

- **Usage Statistics**
- 4个月内， A total of 13,630 people played the game during this time, generating 1,271,451 labels for 293,760 different images. Over 80% of the people played on more than one occasion (i.e., more than 80% of the people played on multiple dates). Furthermore, 33 people played more than 1,000 games (this is over 50 hours of playing!).
- The average number of labels collected per minute by a pair of individuals is 3.89 (std. dev. = 0.69).

- **Quality of the Labels**
- Car dog, man, woman, stamp, Witherspoon (as in “Reese Witherspoon”), smiling, Alias (the TV show), cartoon, and green.



- Please type the six individual words that you feel best describe the contents of this image. Type one word per line below; words should be less than 13 characters

- Current GWAP implementations rely on player collaborations to gather desired information
  - Players in the ESP Game are given the same image, and the descriptions which they both agree become the image labels
  - Players will be rewarded, e.g., gaining game points, for achieving consensus opinions



# Known Issue of ESP Game

- Some known issues in this kind of collaborative mechanism.
  - First, players can benefit from cheating by forming coalitions
  - Second, players tend to give easier and more generic descriptions

# Solution-- **KissKissBan**

- player-level cheating proof mechanism
  - Introduces an additional player, the blocker, whose objective is to stop the matching from happening

# Rule of Game

- In the beginning of each round, the blocker has 7 seconds to provide blocked word list, which is the list of words he/she thinks couples might match on. These blocked words are not visible to the couples.
- After the 7 seconds of entering blocked words, the couples have 30 seconds to match with each other.
- The game time will decrease by 5 seconds if any couple types the blocked word, i.e., being blocked. Also, agreeing on the blocked word does not count as matching.
- The couples win the round if they successfully match with each other within the time limit, otherwise the blocker wins.
- Players switch roles every 5 rounds in 15 rounds of the game.

# Role

- Blocker:
- Each player will play the blocker for 5 rounds in the 15 rounds of the game.
- Though the blocker only has 7 seconds to act in each round, he/she is able to see every word the couples are typing during the game.
- Monitoring the actions of the couples not only makes the waiting process fun, but provides the blocker an opportunity to stop the couples from achieving some unified strategy.
  - For example, the blocker could give “a” as the blocked word if he/she finds the couples try to match on “a” in every round.

- Couple:
- The objective of the couples is the same as the players in the ESP Game: to guess what the partner is typing.
- However, unlike the players in the ESP Game, the couples in KKB cannot see what the blocked words are.
- Therefore, the couples are encouraged to guess harder words to avoid guessing the word in the blocked words list.

# Kiss KISSBAN

Time

0:06

Score

600



Couple1

car  
gun  
fire  
fight

Couple2

lights  
light  
shadow  
war  
army

Curses

yellow  
night



# Kiss KISS BAN

Time

0:22

Score

-100



Correct

Multiple

Curses

1 guess

woman

2 curses

Being Blocked: woman

Game Time - 5

Send

Hint: Guess harder words may prevent being blocked!

# Incentive Structure

- Zero sum game
  - the blocker loses 200 points and each couple gains 100 points when the couples win;
  - the blocker gains 200 points and each couple loses 100 points when the couples lose.
- Available time period for entering words
  - In order to increase the possibility of blocking, the blocker will have to type as many words as possible in the short period of time, i.e., 7 seconds.



# Collected Labels

- two kinds of labels,
  - matching label
  - blocking label
- blocking label is the matching between the couple and the blocker.

# 与TabooWords区别

## 区别

- taboo words are generated from the statistics of past play history, but the blocked words are entered by the blocker in each play.
- Second, the taboo words are visible by players but the blocked words are not.

## 优点

- provides more uncertainties
- couples would be motivated to provide more diverse labeling
- Not biasing player behavior.

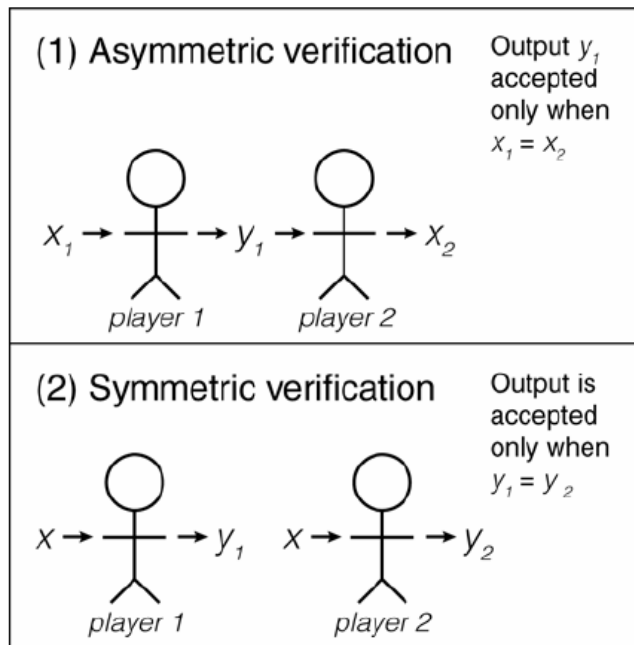


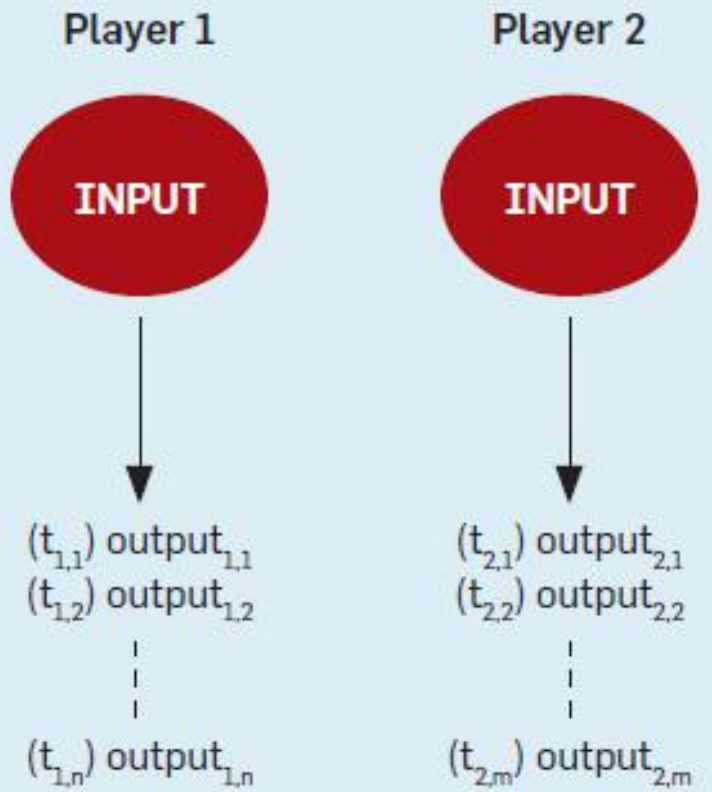
TABLE II  
EXAMPLES OF SOCIAL GAMES

Game Structure	Verification Method	Game Mechanism	Player Requirement		Examples
			Num of Player	Game Play	
Output-agreement	Symmetric	Collaborative	2	Synchronous	ESP, Matchi, Squigl, OntoGame
		Hybrid	Multi-players	Synchronous	Common Consensus, Social Heroes
		Hybrid	Multi-players	Asynchronous	Gopher Game
Input-agreement	Symmetric	Collaborative	2	Synchronous	TagATune
		Hybrid	N/A	N/A	N/A
Inversion-problem	Asymmetric	Collaborative	1 or 2	Synchronous	Peekaboom, Verbosity
		Competitive	2	Asynchronous	Dogear, CyPRESS, CARS
		Hybrid	1 or Multi-players	Synchronous	Phetch
Output-optimization	Symmetric	Collaborative	2	Synchronous	Restaurant Game
		Competitive	N/A	N/A	N/A
		Hybrid	Multi-players	Synchronous	Diplomacy
	Asymmetric	Collaborative	N/A	N/A	N/A
		Competitive	N/A	N/A	N/A
		Hybrid	N/A	N/A	N/A

# Output agreement

- ESP game
- Squigl
- PictureThis
- **Thumbs-Up**

**Figure 1: In this output-agreement game, players are given the same input and must agree on an appropriate output.**



Players win if/when  $\text{output}_{1,i} = \text{output}_{2,j}$

**Figure 2: In this output-agreement game, the partners are agreeing on a label.**



- Squigl

- a game for gathering segmentation data for images in which two players are shown the same image and an associated label, then are asked to draw an outline around the object in the image with that label.
- Points are awarded based on how much the two outlines of the object overlap

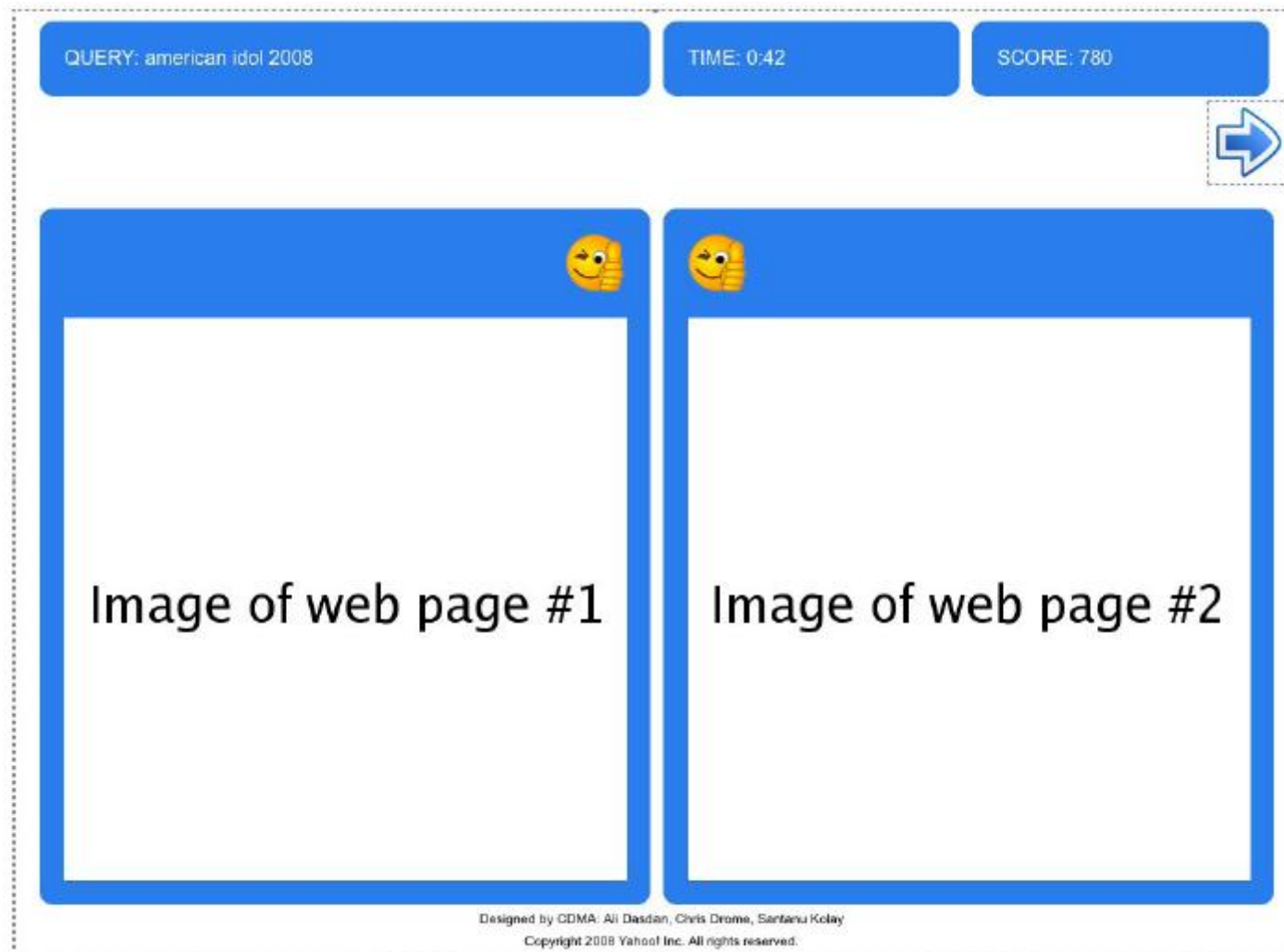
- PictureThis

- Players are shown a label and a list of images and asked to select the image that is the most relevant to that label. Players are again rewarded if their selections match.

- Ontogame and Ontotube
  - players are given various types of input objects (e.g., Wikipedia excerpts, YouTube videos, eBay auctions) and an ontology, then asked to annotate the input object using the given ontology.



# Thumbs-Up:



**Figure 1: Screenshot of Thumbs-Up.**

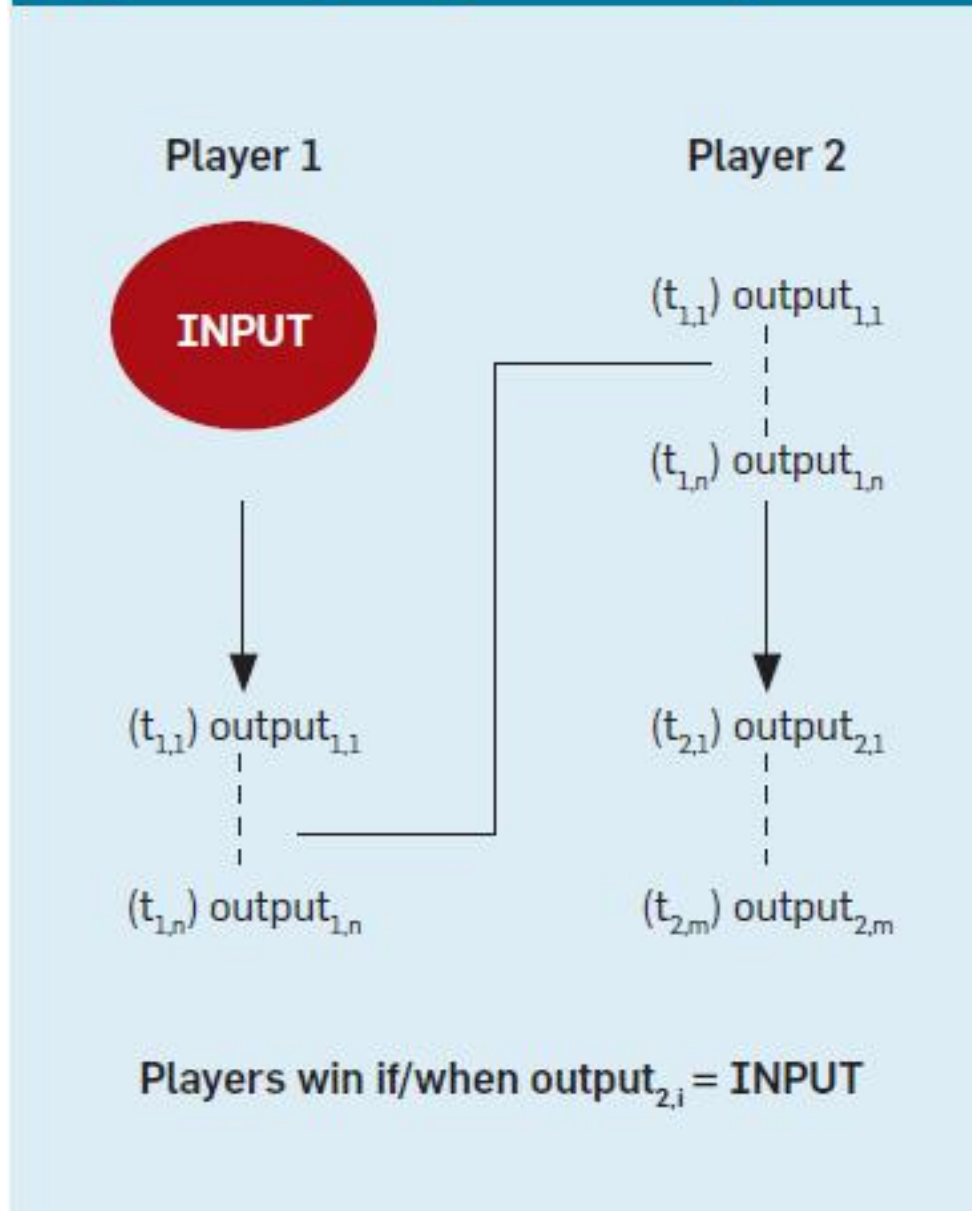
# Game rules.

- A player logs in and is randomly matched with another player. Both players are shown the same input query and images of two web pages deemed relevant to the query. To increase their scores, the players must agree on the same page as more relevant.

# Game features.

- several well-known game features to make Thumbs-Up challenging and fun, which include:
  - a time limit (60 s),
  - score keeping (60 per successful match),
  - daily and all time high score lists, and
  - randomness in selecting partners, queries, and images. Features

**Figure 3: In this inversion-problem game, given an input, Player 1 produces an output, and Player 2 guesses the input.**



# Peekaboom

★ YOU AND A RANDOM PARTNER TAKE TURNS PEEKING AND BOOMING ★

PEEK : GUESS WHAT YOUR PARTNER IS REVEALING

TIME LEFT  
2:23

SCORE  
300

peek  
HINT

GUESSES

fur	100
cat	100
dog	100
sheep	100
horse	100

GUESS HERE

PASS

HINTS HELP YOU GUESS

PASS FOR DIFFICULT IMAGES

BOOM : REVEAL PARTS OF THE IMAGE TO YOUR PARTNER

TIME LEFT  
2:23

SCORE  
300

COW

boom  
CLICK THE IMAGE TO REVEAL AREAS TO YOUR PARTNER  
RIGHT CLICK TO PING AN AREA OF THE IMAGE

GUESSES

fur	100
cat	100
dog	100
sheep	100
horse	100

HINTS

PASS

GIVE HINTS IF NECESSARY

TELL YOUR PARTNER IF A GUESS IS HOT OR COLD

# Peekaboom

- Improves on the data collected by the ESP Game and, for each object in the image, outputs precise location information

as well as other information useful for training computer vision algorithms

- The goal of the game is for Boom to reveal parts of the image to Peek, so that Peek can guess the associated word.
- Boom reveals circular areas of the image by clicking. A click reveals an area with a 20-pixel radius.
- Peek, on the other hand, can enter guesses of what Boom's word is.
- Boom can see Peek's guesses and can indicate whether they are hot or cold.
- When Peek correctly guesses the word, the players get points and switch roles

# elephant, trunk, tusk, ear?

- Pings Component





- **Hints**
- give hints to Peek about how the word relates to the image



- **Game Points and the Bonus Round**

# “bite-size” milestones

TIME LEFT  
2:19

  
**bonus**



CLICK ON THIS  
OBJECT WITHIN  
THE IMAGE.

**car**



**Bonus Round!**

# Collecting Image Metadata

- Input: an image-word pair
- Peekaboom collects the following information:
  - **How the word relates to the image**
  - **Pixels necessary to guess the word**
  - **The pixels inside the object, animal, or person**
  - **The most salient aspects of the objects in the image**
  - **Elimination of poor image-word pairs.**

# Cheating

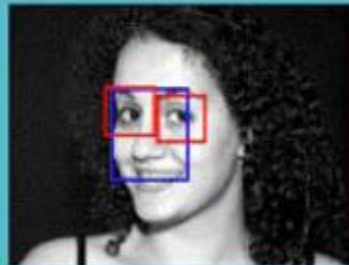
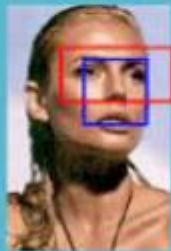
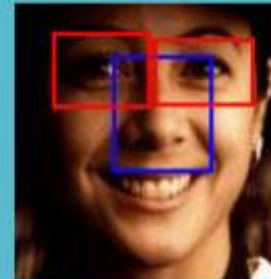
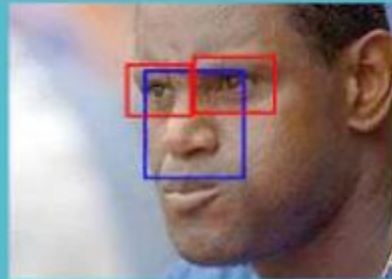
- **The player queue**
- **IP address checks**
- **Seed images**
- **Limited freedom to enter guesses**
- **Aggregating data from multiple players.**

# 用途

- *Improving Image-Search Results*
- *Object Bounding-Boxes*

Peekaboom

49 matches; 891 near matches for **eyes nose**



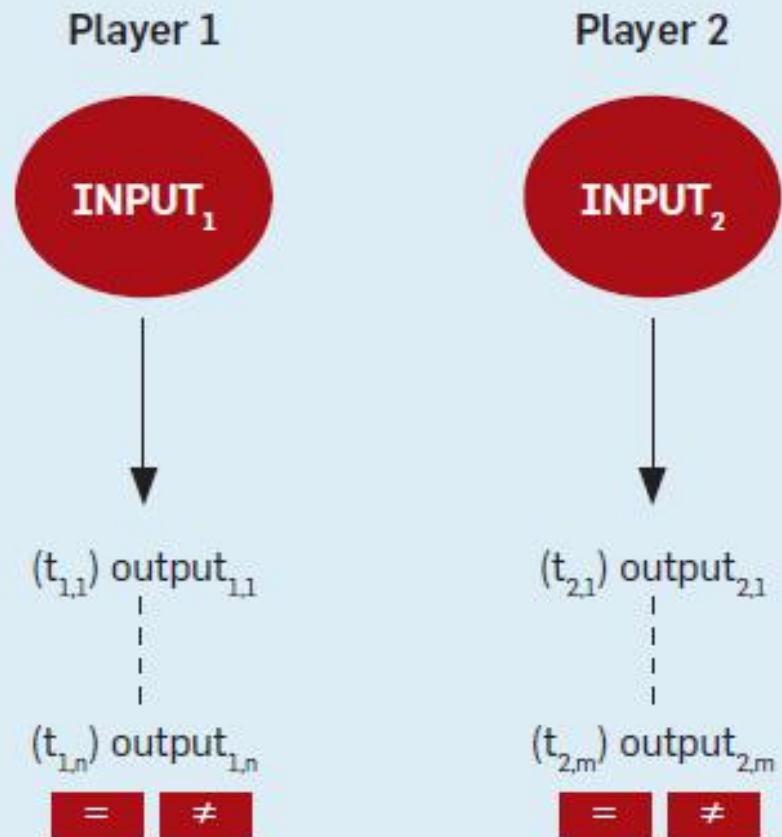
- arrows or lines pointing to the objects





# Input Agreement

**Figure 4: In this input-agreement game, players must determine whether they have been given the same input.**



Win if players guess whether  $\text{INPUT}_1 = \text{INPUT}_2$

# Input Agreement-TagATune

- social tags
  - 28% of Internet users have tagged photos, news stories, or blog posts online
  - Flickr.com (photo sharing), Last.fm (music sharing) and YouTube.com (video sharing)

# known issues

- Only the popular items are typically tagged, leaving a large proportion of the multimedia objects on the Web untagged
- For multimedia objects with a time component, such as sound, music, and video clips, social tags found online often describe the object as a whole, making it difficult to link tags with specific content elements

# ESP Game?

- output-agreement mechanism that works so well in many games failed to work for collecting data about sound clips in the TagATune prototype



# Reason

- it can be very difficult for two players to agree on a description
  - such as “temperature” (e.g., chilly, warm), mood (e.g., dark, angry, mysterious), or the image it evokes (e.g., busy streets, festival), as well as categorizations that have no clearly defined boundaries (e.g., acid-jazz, jazz-funk, smooth jazz).

# New Game

- two players are shown either the same object or different objects and each is asked to type a description of their given object
- Unlike output-agreement games, where all communication is forbidden, all of the players' descriptions are revealed to each other
- Based on these descriptions, the players must decide whether they have been given the same object.

# TagATune interface

The screenshot displays the Tag a Tune game interface. On the left is a 'Most Points Today' leaderboard with 10 entries. The main game area shows a score of 80 and a timer of 1:41. A 'Describe the tune ...' section includes a play button and a progress bar. A 'Listening to the same tune?' section has 'same' and 'different' buttons. A central notification bubble reads 'Correct 60 points' with 'You' and 'Partner' labels and green checkmarks. The 'your descriptions' column lists 'male vocal', 'medieval music', 'quartet', and 'two females'. The 'your partner's descriptions' column lists 'guitar', 'solo', and 'no vocals'. At the bottom, there is a text input field, 'submit', and 'pass' buttons, and a message 'Your partner has chosen.'

Most Points Today	
1	sunshine 173 K
2	quest40992 86 K
3	WhigleyRue 50 K
4	ocean 24 K
5	ScottParade 20 K
6	hain 17 K
7	missy420 16 K
8	adaman 12 K
9	Amro 10 K
10	tonkiddo 9,850

Score: 80  
Timer: 1:41  
Tag a Tune  
Hear Here

Describe the tune ...  
0:10

Listening to the same tune?  
same different 1 in a row

your descriptions  
male vocal  
medieval music  
quartet  
two females

your partner's descriptions  
guitar  
solo  
no vocals

You Correct Partner  
60 points

submit pass

Your partner has chosen.



For manuscript's sake, I have written some notes  
at the end of the book, for your reference. Overall your writing style  
is a bit too heavy. You do not have good penmanship,  
but they get lost amidst the words.  
S.



- 通过游戏判断文本中表达的情绪
- 游戏: Guesstiment
  - Guesstiment is a two-player asynchronous game
  - Aims at annotating a large corpus of text documents, similar to the goal of the ESP game
  - 与ESP的不同
    - Rules
    - Asynchronous approach

# two players

- Suggester
  - starts each round
  - Be given the whole text of a review document and he/she is supposed to:
    - Decide whether the whole text is positive or negative, i.e. the author is praising about a subject or criticising it.
    - Select a single word (or a sequence of words, as short as possible) which best describes the polarity (positive or negative) he has selected in part (1).
      - For example, when the negative polarity is chosen, the word "terrible" would be a good choice for the representative word (provided that it is present in the text).

- Guesser

- be given only the word (or word sequence) suggested by the Suggester
- guess polarity of the whole text just based on that single word.

# Suggester Role

Startup Suggest Guess Help Top Scores

The Article

Worst hotel Experience We booked our stay for two nights via a special on their website. I am a Starwood Member. I am not the type to complain about these sort of things. We had to change rooms twice. the first room was large but the bed sank in the center in a cartoonish way when someone sat on it Also, there were no soap or bathmats in the bathroom and there was a layer of dust over everything, as if the room had not been dusted in weeks. the white comforter on the bed was DIRTY. At first we attempted to get clean sheets and towels, but after two calls, all we received were 2 more bars of soap and some ratty bath mats. We switched rooms which meant bringing all of our luggage down to the

Forbidden Words

for  
no  
of  
rooms  
banged  
not  
view  
bar  
breakfast

Which word describes the polarity (or at least its surrounding text) best?

And what do you think polarity is?

Game information received

Total Score: 0 Last Round Score: -

# Guesser Role

Startup Suggest **Guess** Help Top Scores

If the following has been selected as a representative for the polarity of whole text...

filthy

And what do you think polarity is?

▼

Game information received

Total Score: 2 Last Round Score: 2