LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Abschlussvortrag Bartholomäus Steinmayr

Designing Image Labeling Games for More Informative Tags

(Diplomarbeit)

Betreuer: Fabian Kneißl, Christoph Wieser
Verantwortl. Hochschullehrer: Prof. Dr. François Bry

Overview

- Problem recapitulation
- Solution: “Karido” implementation
- Evaluation and results
- Conclusion and future work

Games With A Purpose

- Some problems are hard for computers, but easy for people:
  - Text and speech recognition
  - Translation of natural language
  - Semantic image analysis
- Games With A Purpose
  - Utilize human cognition for solving these problems
  - Should be fun to play
  - Should collect valid data

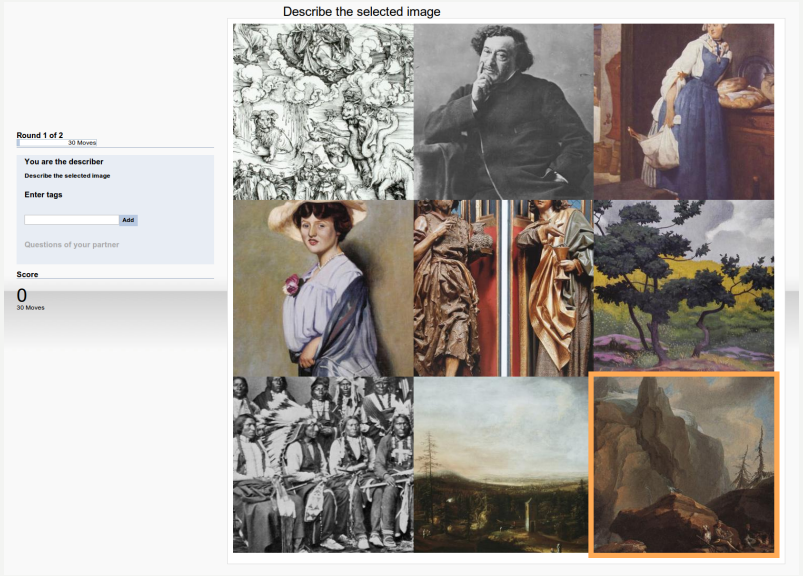
Motivation

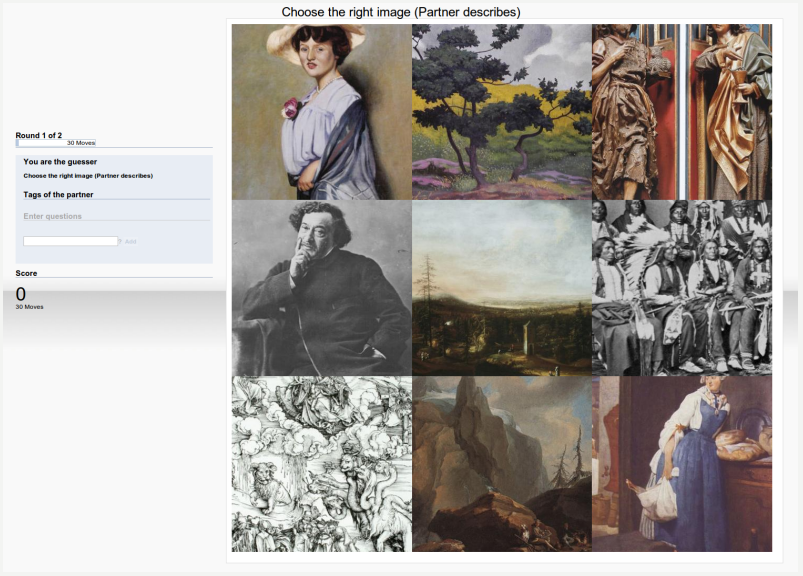
- Intention of image labeling games:
  - Acquiring both general and specific tags
- Existing game designs favor general Tags
  - Proven using game theory by Jain and Parkes
- Current solution:
  - Taboo words
  - Demonstrated to be insufficient by Weber et al.

Proposed Game

“Karido” key features:

- Asymmetric game design
- One player describes, other player guesses
- Guesser can ask questions
- Image selection by similarity





Implementation

- Game modes:
  - Time-limited mode
    - 90 second limit per round
    - Focus on high number of collected Tags
    - Might be considered stressful
  - Turn-limited mode
    - 30 actions per round
    - Actions: Description, Question, Wrong Guess
    - Players take turns
    - Focus on quality of collected Tags

Tag Verification

- ESP Game:
  - Matches between players are used to verify Tags
- Karido:
  - Random guessing is successful with high probability
  - “Matches” between players not sufficient for verification
- Solution:
  - Score penalty for wrong guesses
  - => malicious users could still introduce wrong data

Tag Verification

- Non-binary verification method
  - Real-valued verification score for each Tag/Image pair
  - Score is increased if Guesser succeeds
  - Amount of increase depends on number of wrong guesses:
    - More than 30% wrong: No increase
    - Right on first try: Increase of 1.0
    - In-between: Linear interpolation
  - Amount of increase depends on number and position of Tag:
    - 0.5 base weight for all Tags
    - Linearly increasing additional weight

Simulated Player

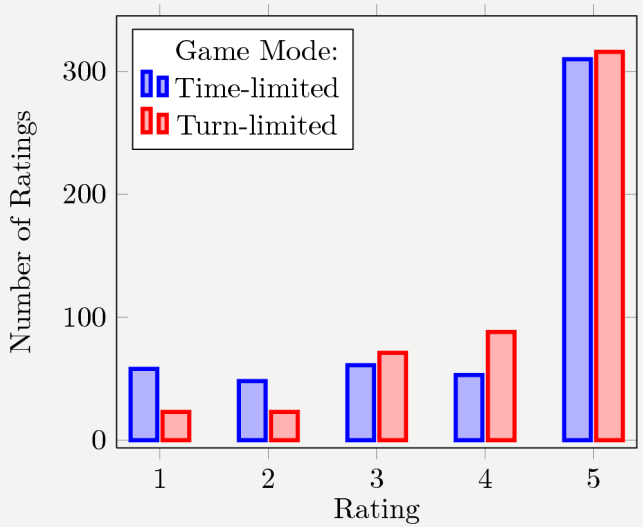
- Simulation of Describer:
  - Does not have to respond to human player
  - Replaying of previously recorded rounds
  - Fallback solution: Random selection of previous Tags
  - Questions answered based on existing Tags

Simulated Player

- Simulation of Guesser:
  - Has to respond to input of human player
  - Heuristic: Image selection based on existing Tags
  - Problem: New Tags
    - Should be rewarded
    - May be erroneous
  - Problem: Questions
    - Should be specific
    - Should not be obscure

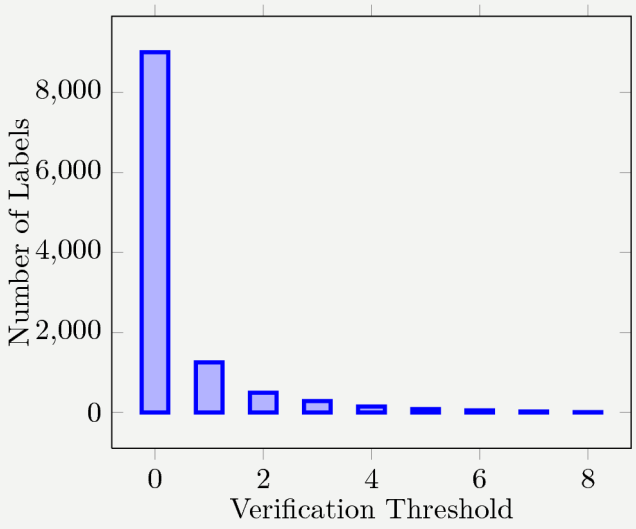
Evaluation

- New game was made public April 13.
- Six week trial period
- Results:
  - 3542 played rounds, 1939 played sessions, 1051 ratings
  - High player satisfaction (4.0/5 for time-limit, 4.2/5 for turn limit):



Results

- Tagging:
  - 6933 unique Tags (16635 for Artigo)
  - 2.0 unique Tags per round (1.2 for Artigo)
    - First three weeks: 2.1/1.6
  - 4.6 Taggings per round (11.0 for Artigo)
  - Small number of verified tags:



Conclusion / Future Work

- Contribution
  - Creation of a novel image labeling game
  - Creation of a framework for multi-player interaction
  - Preliminary evaluation suggests:
    - High player acceptance
    - High number of unique Tags
- Future work
  - Large-scale evaluation
  - Collection of compound Tags
  - Use of Karido for other media
  - Implementation of mobile game

Thank you for your time!

I am looking forward to your questions and suggestions.