System and Method of Automating Psychophysics Experiments

Yue Qiao
IBM Printing Systems Division
Boulder, Colorado

Abstract

This paper describes the system and method of automating psychophysics experiments. The system was designed consisting five components, representing different functionalities and workflow stages within a psychophysics experiment, experiment design and setup, observer information input, experiment, data analysis, and report. Seven psychophysical evaluation methods were automated in the system: paired comparison, rank order, ratio scales, rating scales, category scaling, multidimensional scaling, and probit analysis.

Introduction

Psychophysics is the scientific study of the relationships between the physical measurements of stimuli and the sensations and perceptions that those stimuli evoke.[1] It is the most reliable method for evaluating image quality of various imaging systems. However, running psychophysical experiment is time consuming, labor intensive, and often error-prone. It requires careful planning and painstaking efforts. For each experiment, the experimenter needs to randomly sort sample orders for each observer, register each observer's information, including name, age, gender, experience level, etc., record each observer's responses, analyze data by hand, and write a report. Thus, it would be of great benefit if the process of psychophysics experiments, from experiment setup through report generation, could be automated in one application system. The IBM Automated Psychophysics Tool (APT) was designed and implemented to accomplish this goal.

Overview of the System Design

There are five components in the Automated Psychophysics Tool (APT). Each component captures different objectives and workflow phases during a psychophysics experiment. They are experiment design and setup, observer information registration, experiment, data analysis, and report. Figure 1 demonstrates the architecture design and workflow of APT.

The experimenter initiates the experiment in the Setup window, and the samples are generated according to the experimental methods chosen for each observer. The random order of the samples for each observer is forwarded to the Experiment window. Once the experiment is set up, the observer enters the information in the Observer Information Window, and starts the experiment following the on-line instruction appeared on the top of the Experiment window. The observer information and the observer's responses that observer entered are recorded and analyzed in the Data Analysis component. The results of the evaluation are displayed in both text and graphics. The report of the experiment is generated automatically once the experimenter decides to end the experiment.

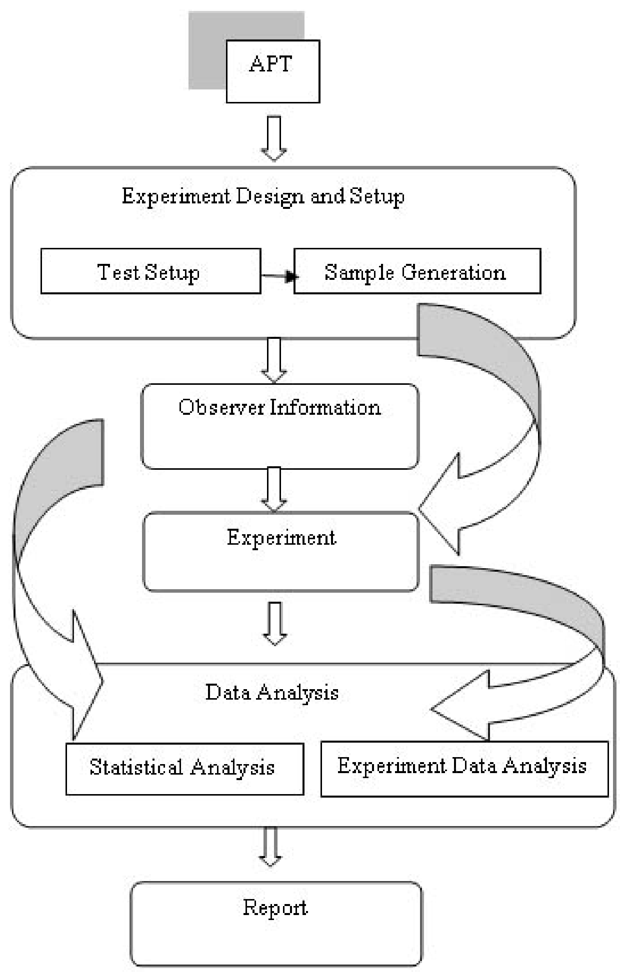

Figure 1. APT design and workflow.

System Workflow

Experiment Design and Setup

This component enables the experimenter to design, setup, and initiate the test. There are two windows in this component: experiment setup and sample generation. Seven fields are listed in the setup window:

- Experiment Name
- Number of Test Sets
- Viewing Condition
- Test Set Names and Description
- Experimental Methods:
  - paired comparison,
  - rank order,
  - ratio scales,
  - rating scales,
  - category scaling,
  - multidimensional scaling,
  - probit analysis.
- Comparison Objects
- Test Samples

The system allows multiple experiment sets to be combined into one experiment. Each experiment set is normally targeted at a different task. This enables the experimenter to combine several small tasks at once for one group of observers, so the resources (observers' time) can be utilized more efficiently. The system also allows multiple evaluation methods and different test samples for each experiment set.

The experimenter then specifies the viewing condition for the experiment, and defines the name, the description, the experiment methods, the comparison objects, and test samples for each experiment set. All these information are recorded and analyzed in the report.

The test samples can be downloaded to the system for displaying in the Experiment window. The images of the test samples are synchronized with the order of the test samples.

After the experiment is setup, the system generates samples according to the experimental method, and randomizes the sample order for each observer.

Figures 2 and 3 demonstrate experiment set up and sample generation.

Observer Information

The following fields are required for the Observer Information:

- Observer Name
- Age
- Gender
- Company
- Occupation
- Observer Experience Level

The statistical information of observers will be analyzed by the Data Analysis component, and will be recorded in the report.

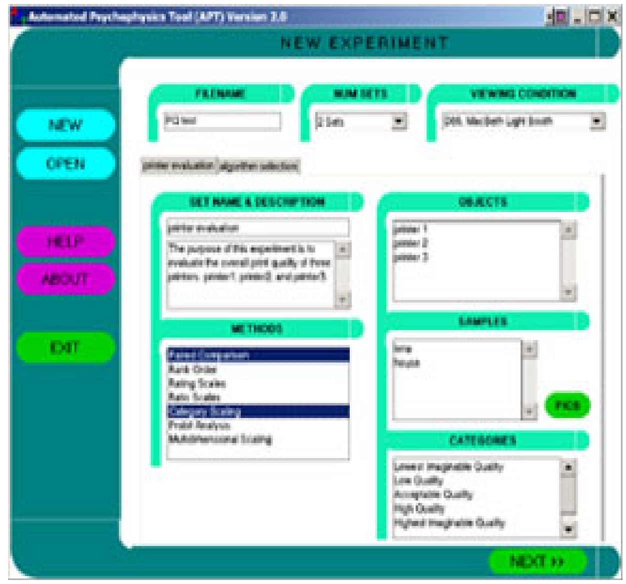

Figure 2. Example of experiment setup.

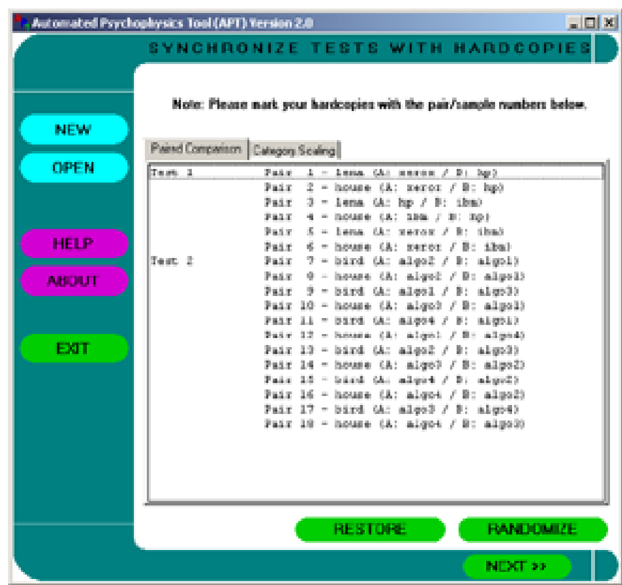

Figure 3. Example of sample generation.

Figure 4 is an example of Observer Information window.

Experiment

The Experiment component was designed to record responses for each observer.

At the top of the experiment window, the instruction is provided for the selected experiment method.

The image of the sample is shown at the bottom of the Experiment window when the observer is making the judgment over that sample. This will ensure the observer judging the samples in the correct order. The randomized order of the test samples are created from the sample generation window, and synchronized with the original sample set.

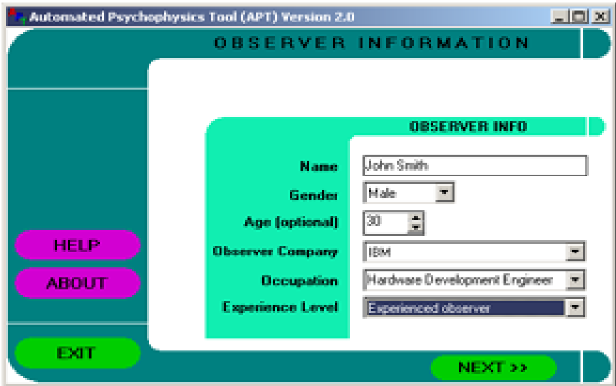

Figure 4. Example observer information window.

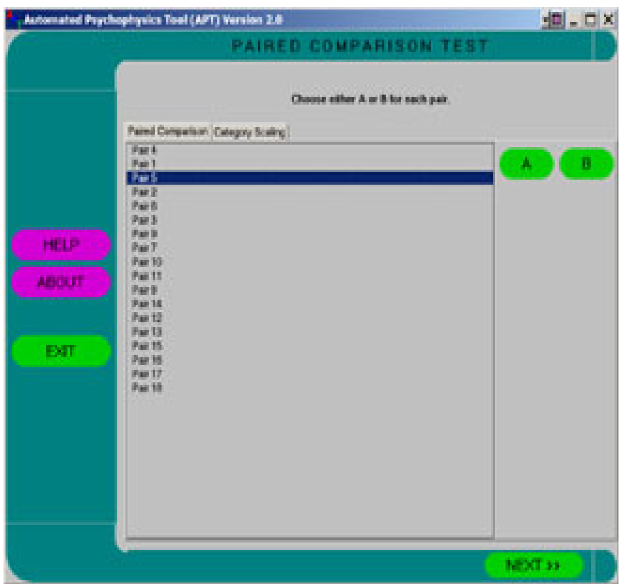

Figure 5. Example observer experiment window for paired comparison method.

Figure 5 demonstrates paired comparison experiment.

Data Analysis

The system performs both statistical analysis of the observers, and data analysis of observers' judgments according to the psychophysics evaluation methods. Both rank order and paired comparison methods were calculated based on the law of comparative judgments. Category scaling method was obtained by the law of categorical judgments. Probit Analysis uses the assumption of cumulative normal distribution. Method of triadic combination: "most alike" and "most different" was used in MDS. Both metric MDS and non-metric MDS were applied for the data analysis.

The set of data included in the analysis process is configurable. By default, the system calculates the overall results for the combination of all the samples. It can also calculate the experimental results for any combinations of test samples by selecting and deselecting samples in the Data Analysis window.

The results of the statistical analysis will be included in the report as part of experiment description. The results of the experimental data analysis for each method are displayed in both text and graphic formats.

Figure 6 demonstrates data analysis for paired comparison experiment.

Figure 6. Example of the data analysis window for paired comparison method.

Report

The report is automatically generated once the experimenter ends the experiment. The report includes following features:

- Introduction: experiment description
- Experiment: experimental methods, test samples selection, viewing condition, and statistical analysis of observers
- Results: by default, the overall test results in text and graphics for all samples are included in the report. However, the test results for the specific combination of test samples can also be added to the report.

Summary

The Automated Psychophysics Tool (APT) was designed and implemented to automate the complex, tedious and often error-prone psychophysics experimental process. With five components representing different workflow phases of psychophysics experiments, and seven psychophysical evaluation methods implemented, APT significantly shortened the length of the psychophysics process, and improved efficiency, accuracy, effectiveness, and productivity.

References

1. Mark Fairchild, Vision and Psychophysics, RIT class material.

Biography

Yue Qiao is an advisory color scientist at IBM Printing System Division, Boulder, Colorado. She received her master degrees in physics and imaging science from John Carroll University and Rochester Institute of Technology, respectively. Since joining IBM in 1996, she has been working on and leading numerous color and print quality related projects, including color management, color halftoning, and print quality evaluations. She is currently pursuing her Ph.D. degree in mathematics at Colorado State University.