RTN2-2001-00107

Perception for Recognition and Action (PRA)

http://pra.psy.gla.ac.uk

Glasgow

London

Toulouse

Marseille

Giessen

Tübingen

Leuven

Rotterdam

P ***erception for***
R ***ecognition and***
A ***ction***

Abstract

One of the greatest scientific challenges of the new millennium is to understand the brain processes that lead to behaviour. Organised around three large-scale projects, the scientific aim of this research proposal is to examine the basic mechanisms of visual perception for the purposes of both scene recognition and action planning. Applications include safer driving environments and new therapies for brain-injured patients. The network integrates the multi-disciplinary expertise of eight pre-eminent European research teams in cognitive neuroscience and builds on existing bilateral collaborations. The network will be an outstanding springboard for the future scientific career of ten young researchers and help Europe regain its leading role in visual perception research. The training programme integrates local individual training, workshops, training courses and exchanges of young researchers across different teams. Through active scientific collaborations and an intensive training program, we will establish a European-wide scientific network that will have an impact well beyond the three years of the proposal.

1. Members of the Network

1. University of Glasgow [UGlasgow], established in the United Kingdom (principal contractor);
2. Royal Holloway and Bedford New College, University of London [RoyalHolloway], established in the United Kingdom;
3. Centre National de la Recherche Scientifique, Délégation Midi-Pyrénées [CNRSToulouse], established in France;
4. Centre National de la Recherche Scientifique, Délégation Provence [CNRSMarseille], established in France;
5. Justus-Liebig-Universität Giessen [UGiessen], established in Germany;
6. Max-Planck-Institute for Biological Cybernetics, Tuebingen [MPITuebingen], established in Germany;
7. Katholike Universiteit Leuven [KULeuven], established in Belgium;
8. Erasmus University [UErasmus], established in the Netherlands.

2. Brief Description of the Work

The purpose of the present research proposal is to understand the fundamental mechanisms underlying visual perception for the purposes of scene recognition and action planning. We propose to address this objective using three complementary large-scale projects. First, we shall study perceptual aspects of recognition by analysing chromatic and achromatic cues for scene recognition, how human observers attend to diagnostic recognition information, and how these different visual attributes are bound together to solve visual tasks. Second, we shall focus on perceptual aspects of action by studying the visual segmentation problem, the usage of three-dimensional cues for action, and the cerebral pathways underlying these visuo-motor tasks. Finally, we shall study the interaction between recognition and action by focusing on the high-level representations for action planning, the eye movements used during scene exploration, and the imitation of human movement. These three large-scale projects will each involve three working groups whose role will be to supervise two or three milestone experiments. Each working group will use a variety of well-established and newer experimental paradigms borrowed from psychology, motor control and brain imaging. For instance, the perception of motion transparency will be addressed both from a psychophysical point of view to determine the efficiency of human observers in this task, but also with the help of fMRI studies to reveal the brain structures involved in this task. Our ultimate goal is to increase our understanding of perception for recognition and action through active collaboration and shared expertise of eight leading laboratories in Europe in these fields. We therefore anticipate major scientific breakthroughs from this unique joint effort in the following directions: a better understanding of how high-level, object-related visual cognition interacts with low-level, image-related processing to offer a new theoretical framework for artificial vision research; a better description of motor behaviour in terms of sensorimotor transformation where complexity of visual processing is fully taken into consideration; a better understanding of cortical dynamics involved in both perception and action.

3. Research Topic

Recognising the face of a famous person or driving a car are just two activities that we perform every day both accurately and effortlessly. In spite of their seeming triviality, the successful execution of these tasks requires the involvement of more than half of the human cortex. Recently developed methods and techniques (such as eye-tracking and functional imaging of the brain) allow us to explore the mechanisms underlying these actions as never before. We expect that our investigation and integration of knowledge will lead to major breakthroughs in our understanding of perception for recognition and action that will not only give new insights for cognitive neuroscientists but also help our community in everyday situations that involve an interaction between perception and action. These everyday situations include activities such as driving cars, operating machinery, and participating in sport, and can be highly taxing for the growing elderly population of Europe. Moreover, we believe we will increase our understanding of brain disorders such as visual agnosia and apraxia, thereby paving the way for the development of new rehabilitation therapies.

In spite of our increasing knowledge of the early stages of visual processing, we are still largely ignorant about how visual information is further processed for higher-level tasks such as face recognition or driving. This ignorance is exacerbated by an increasing division of perceptual research into specialised fields. Vision research is currently divided between low-level and high-level tasks with little interaction between these two fields. Some theoretical limitations result from this dichotomy. Although spatio-temporal properties of low-level visual processing are thought to constrain higher-level tasks such as face recognition, very little is known about how recognition processes affect very early, filtering stages of visual pathways through top-down influence. Therefore, most computational models ignore top-down feedback for active selection of the most important visual attributes for a given task. One strength of the present network is the use of various methodologies ranging broadly from low-level psychophysics to high-level cognitive psychology.

Inter-disciplinary methods will be used to tackle one of the most influential theories in cognitive neuroscience. According to this theory, the perceiving and acting brain work largely independently. Although there is an extensive literature claiming that perception and action are tightly coupled in living organisms, most current research tackles one or the other aspect of the coupling, ignoring its complementary part. For instance, face recognition and eye-movement studies are two areas of research that have become fairly autonomous despite the evidence that observers' saccadic scan pathways of face images are highly organised. This division of expertise makes the integration of knowledge difficult, blocks the development of adequate experimental paradigms to investigate perception-action coupling and produces over-simplistic theories claiming that there are two, largely independent visual subsystems: a visuo-motor and a visuo-perceptual. In addition, training of students is made more difficult as scientific development calls for highly specialised researchers but also requires more inter-disciplinarity. With the help of multi-disciplinary expertise, the network will challenge the dominant view that perception and action are independent mechanisms.

The present proposal benefits from the collaboration of eight research groups with complementary expertise. This expertise is centred on three major themes: (1) low-level perceptual attributes; (2) high-level recognition abilities; and (3) high-level action abilities. The goal of our research proposal is to link together these themes. In particular, we shall reveal how perceptual representations vary between recognition and action tasks. We will bridge the gap between low-level and high-level perceptuo-motor tasks by bringing together research teams whose expertise covers at least two of the above themes. In summary, our proposal will cover the following three Large-Scale Projects (LSPs): (A) perceptual attributes for recognition; (B) perceptual attributes for action; and (C) interaction between recognition and action.

4. Research Method

The study of visual perception has outstanding potential as an inter-disciplinary enterprise. Results from anatomy, physiology, psychophysics, experimental psychology and computer science can be combined to further our understanding of the human visual system. Our research methods will range from detailed psychophysical threshold measurements to functional brain imaging. As detailed below, the different teams in the network have extensive and complementary knowledge about these state-of-the-art tools. The adequacy of the methods is substantiated by:

- their high accuracy and robustness in measuring human performance,
- the fact that they are non-constraining and non-invasive for the participants, and
- their flexibility and ease of use that will facilitate the transfer of expertise between teams.

While some of these research methods are relatively new, extreme care will be taken to follow the ethical regulations and guidelines of international and institutional authorities. In particular, new functional imaging techniques involve placing the participant within an intense electro-magnetic field to record their brain activity, but independent research has failed to demonstrate any risk associated with this technique. All experiments will be run only with the full prior consent of the participants and after approval by the local Ethic Committees at each of the participating sites.

4.1. Experimental psychology and visual psychophysics

All participating teams have access to state-of-the-art equipment for investigating visual perception at different levels. An original aspect of this project is that the equipment of each member team complements the others', thus offering students the opportunity to complete projects with virtually no technical limitation. Depending on the type of research done in each team, complementary techniques are available for low-level vision (VSG graphics cards, MATLAB-driven PCs and SGI workstations), together with stereographic displays (Wheatstone stereoscope, shutter-goggles) as well as for mid- and high-level vision (Graphic SGI server, virtual reality displays, large fields displays). Customised software has been developed within each team for several years and will be exchanged among collaborating teams. Classical psychophysical procedures will be used routinely (forced-choice paradigms, adaptive methods for threshold estimation, threshold tracking…). In some cases, event-related potentials (ERPs) will be measured using recording techniques available in several of the member teams (UGlasgow, CNRSMarseille and CNRSToulouse). Experimental psychology and visual psychophysics will be used by all working groups and will constitute the common language of the network.

4.2. Eye movements recordings

Eye movements can be recorded with great accuracy in at least five of the eight member teams. Non-invasive techniques (using Purkinje trackers and other video eye trackers such as EyeLink) have been developed over the last ten years in KULeuven, CNRSMarseille, UErasmus, UMagdeburg and RoyalHolloway. The more stringent, but also much more accurate, search coil technique is also available in the CNRSMarseille and UErasmus teams for 2D tracking eye movements as well as for vergence recordings. Moreover, vergence movements can be studied in KULeuven where two dual-Purkinje-image trackers can be used (one for each eye). Customised software is available to these teams and technical support will be provided by them should some of the other teams wish to acquire eye trackers. Eye movement recordings will be used primarily for the second large-scale project (WG-B.1 and WG-B.3), but also by other working groups (WG-A.2 and WG-C.2). In addition, one training course will be organised by UErasmus to review some of the techniques used for eye movement recordings.

4.3. Arm/hand movements recording

Several versatile environments have been set up to record arm and hand movements in real time. MPITübingen, UErasmus and UGlasgow each own an OPTOTRAK optical device to record highly accurate locations of small sensors that can be placed at different joints of the moving limb. RoyalHolloway has a similar system (ELITE). KULeuven owns a motion capture system (ProReflex) to study biological motion perception. MPITübingen has set up a table-top virtual workbench in which 2 force feedback devices (Phantoms) and computer graphics allow to manipulate virtual haptic and visual objects. UGlasgow, RoyalHolloway, MPITübingen and UMagdeburg also own several other movement capture devices such as Flock-of-Bird electro-magnetic tracking devices, a 6 camera synchronised digital video system, and a number of touch screens for pointing tasks. Arm and hand movement recordings will be used mainly for the third large-scale project (WG-B.2, WG-C.1 and WG-C.2). These techniques will be reviewed in the training course organised by UErasmus.

4.4. Functional imaging

Four of the member teams have access to functional imaging techniques (RoyalHolloway, CNRSMarseille, UMagdeburg and KULeuven). CNRSMarseille site has just opened a new 3 Tesla fMRI for imaging studies in Neuroscience. This facility is devoted to research only so that continuous access will be possible for all other member teams. KULeuven has worked on perception and memory using fMRI techniques. UMagdeburg has access to a state-of-the-art imaging centre with several scanners (1.5 Tesla GE, 3 Tesla Bruker head scanner) dedicated to research. CNRSToulouse is expecting to receive a new 3 Tesla fMRI machine as well. Finally, the RoyalHolloway team has long-term experience in fMRI studies and has developed its own analysis tools. The team will have their own 3 Tesla research scanner in London by the end of 2001, and will offer scientific support for the fMRI projects in the network. Functional imaging will be used effectively by most working groups and will be central for WG-B.1 and WG-B.3. Moreover, RoyalHolloway will organise a training course on functional imaging techniques (see Table 9).

4.5. Computational Modelling

Several teams have been involved in the modelling of experimental data. In particular, CNRSToulouse has developed a large-scale neural network simulation system. Currently, they can simulate networks with hundreds of millions of neurones and billions of connections. This network is in the process of being commercialised. The UGlasgow team has worked extensively on statistical modelling of perceptual and motor behaviours, in particular using a Bayesian framework. Computational modelling will be used by several working groups, including WG-A.1, WG-B.1, and WG-B.2.

5. Work Plan

The research activity of the network is centred on three Large-Scale Projects (LSPs). Each LSP is composed of three smaller Working Groups (WGs) that involve a subset of the participating teams in the network (see Figure 1) and that will address the specific research issues described below. Each WG will focus on well-defined series of experiments named milestones.

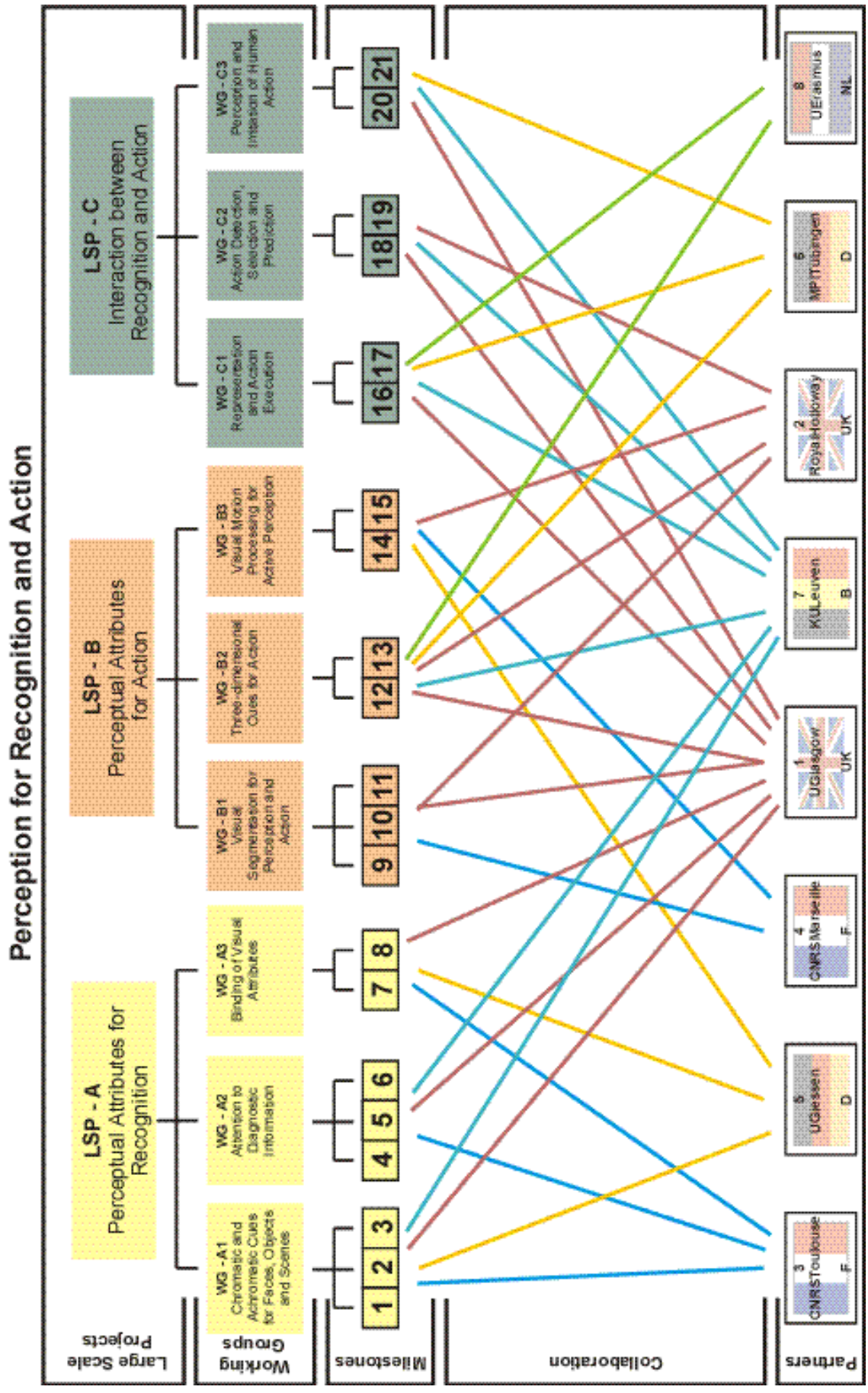
Perception for Recognition and Action
Large Scale Projects
LSP - A
Perceptual Attributes for Recognition
LSP - B
Perceptual Attributes for Action
LSP - C
Interaction between Recognition and Action
Working Groups
WG - A1
Chromatic and Achromatic Cues for Faces, Objects and Scenes
WG - A2
Attension to Diagnostic Information
WG - A3
Binding of Visual Attributes
WG - B1
Visual Segmentation for Perception and Action
WG - B2
Three-dimensional Cues for Action
WG - B3
Visual Motion Processing for Active Perception
WG - C1
Representation and Action Execution
WG - C2
Action Detection, Selection and Prediction
WG - C3
Perception and Imitation of Human Action
Milestones
1 2 3
4 5 6
7 8
9 10 11
12 13
14 15
16 17
18 19
20 21
Collaboration
Partners
3 CNRS Toulouse F
5 UGiessen D
4 CNRS Marseille F
1 UGlasgow UK
7 KULeuven B
2 RoyalHolloway UK
6 MPI Tübingen D
8 UErasmus NL

Figure 1