**Matthew Turk毕业于MIT，最有影响力的研究成果：人脸识别。其和Alex Pentland在1991年发表了”Eigenfaces for Face Recognition”.该论文首次将PCA（Principal Component Analysis主成分分析）引入到人脸识别中，是人脸识别最早期最经典的方法，且被人实现，开源在OpenCV了。主页：[www.cs.ucsb.edu/~mturk/](http://www.cs.ucsb.edu/~mturk/" \t "_blank)**

**特征脸**（Eigenface）是指用于[机器视觉](http://zh.wikipedia.org/wiki/%E6%9C%BA%E5%99%A8%E8%A7%86%E8%A7%89" \o "机器视觉)领域中的[人脸识别](http://zh.wikipedia.org/wiki/%E4%BA%BA%E8%84%B8%E8%AF%86%E5%88%AB)问题的一组[特征向量](http://zh.wikipedia.org/wiki/%E7%89%B9%E5%BE%81%E5%90%91%E9%87%8F)。使用特征脸进行人脸识别的方法首先由Sirovich and Kirby (1987)提出，并由[Matthew Turk](http://zh.wikipedia.org/w/index.php?title=Matthew_Turk&action=edit&redlink=1)和[Alex Pentland](http://zh.wikipedia.org/w/index.php?title=Alex_Pentland&action=edit&redlink=1)用于人脸分类。该方法被认为是第一种有效的人脸识别方法[[来源请求]](http://zh.wikipedia.org/wiki/Wikipedia:%E5%88%97%E6%98%8E%E6%9D%A5%E6%BA%90)。这些特征向量是从高维矢量空间的人脸图像的[协方差矩阵](http://zh.wikipedia.org/wiki/%E5%8D%8F%E6%96%B9%E5%B7%AE%E7%9F%A9%E9%98%B5)计算而来。

**一组特征脸** 可以通过在一大组描述不同人脸的图像上进行[主成分分析](http://zh.wikipedia.org/wiki/%E4%B8%BB%E6%88%90%E5%88%86%E5%88%86%E6%9E%90)（PCA）获得。任意一张人脸图像都可以被认为是这些标准脸的组合。例如，一张人脸图像可能是特征脸1的10%，加上特征脸2的55%，在减去特征脸3的3%。值得注意的是，它不需要太多的特征脸来获得大多数脸的近似组合。另外，由于人脸是通过一系列向量（每个特征脸一个比例值）而不是[数字图像](http://zh.wikipedia.org/wiki/%E6%95%B0%E5%AD%97%E5%9B%BE%E5%83%8F)进行保存，可以节省很多存储空间。

# CUBRICON

系统是布法罗纽约州立大学在高级国防研究工程局的支持下开发的一个基于知识的多媒体系统，可用于空军作战计划等任务中。CUBRICON系统执行的输入模式是语言/键盘或鼠标。

CUBRICON具有以下的功能：

* 结合用户的指点手势和自然语言接受和理解多种媒体的同时输入，并能推理和其伴随语言可能不一致的指点手势的输入；
* 自动选择适合用户表达信息的输出媒体形式，并自动判断信息间的相关性；
* 自动进行窗口操作。

CUBRICON可从三种输入设备接受输入：语音输入设备、键盘和图形显示器上指向目标的点。可同时为三种输出设备提供输出：高分辨彩色图形显示器、单色显示器、语音输出设备

CUBRICON系统接受如下输入媒体：口头自然语言和书写自然语言、指点手势，地图，彩色图形显示器，表和表格。

CUBRICON最重要的特征之一是不仅以多种方式产生输出，而且能够决定用哪种方式以及怎样来利用并结合它们

# QuickSet

笔/声音系统，运行在早期的平板电脑上，曾运用于美国海军陆战队军事训练中

潜在优势：

自然：更加接近人类所熟悉的交流方式

高效：

灵活：用户输入更加灵活

可靠：

* 与只用语音相比，可以识别更短更简单的话语。从而减少不流畅，增加语音识别的鲁棒性，可以针对空间信息提供更高的精度
* 适用于更多环境、用户、任务
* 可以适应连续变化的环境条件

# 输入：

脑机交互，神经活动感知，脑电波

# 输出：

正在逐渐被关注

# 输入到输出：

输入和特定用户行为 到 用户意图的映射并不是简单直接的

举例，简单直接的，手机来电 接听。更多情况下是不直接的，是一个open research question

# 生物感知整合

从出生起，就接受外界的多种输入，身体（大脑）将接受到的多种输入信号，加以整合，得到正确的信息。习以为常，察觉不到

举例：麦格克效应，证明了听觉和视觉是相互影响，并最终影响人类对事物的感知

这是大脑对于来自眼睛和耳朵所提供的矛盾信息的努力猜测，这个理论也证明眼睛（**视觉信息**）对于大脑意识与知觉的影响比其他感觉器官所提供的信息更大

# 多通道整合

主要技术难题

输入流的含义随着环境、任务、用户以及时间的不同而改变

最大的挑战是时间维度的挑战，不同的输入形式有不同的时序约束（在时间上有不同的特点）。一些输入方式提供的信息是及时的、离散的、较稀疏的（比如手势）。而另外一些输入方式提供的信息则是连续的（比如情感）。

一些需要并行整合加以解释，另外一些需要先后连接起来

多通道整合的结果到底是什么？

1. 通过多通道整合从而得到一个多通道结果，该结果可以很好适应时间上离散的计算模式
2. 得到更为复杂的感性活动的表述，该表述可以更好的匹配多通道试图提供的人类交互方式

对多通道间复杂关系的深入研究是有前景的研究途径

# 现实情况：

智能移动设备无处不在，功能强大，适合实现多通道的交互

Kinect \Leap Motion

大量的机会和应用，激励了多通道技术的发展

智能手机的普及使得社会加速进入post-WIMP时代

Post-WIMP

用户更加关注于任务本身，而非如何操纵计算机来完成任务

以下摘自维基百科：

WIMP界面自从提出以来如此流行的的原因是 WIMP在抽象的工作空间，文件，行为上的优势。类似纸张或者文档的文件模式让WIMP界面在普通用户看来非常直观。二维平面显示器上矩形框的表现形式非常适合系统编程人员

然而WIMP界面对于复杂工作的并不友好，比如：计算机辅助设计、大量并行数据的处理、交互类游戏。WIMP通常是[pixel](http://en.wikipedia.org/wiki/Pixel)-hungry（像素饥饿）（屏幕支持），因此有限的屏幕容易让人从手头的工作上分散注意力。因此，用户界面可以更好的封装工作空间，行为和特定的复杂任务对象。需要连续输入信号，展示三维模型或者简单地描述一个没有标准定义不见功能的交互。

基于以上考虑，提出Post-WIMP的概念，并已经得到推广。iPod，ATM屏幕