

โครงร่างโครงการทางวิศวกรรม  
(Senior Project Proposal)

ภาควิชาวิศวกรรมคอมพิวเตอร์  
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

1 กันยายน 2553

ชื่อหัวข้อโครงการ (ภาษาไทย)	ระบบควบคุมวัตถุเสมือน	
ชื่อหัวข้อโครงการ (ภาษาอังกฤษ)	Augmented Reality Object Manipulation System	
ชื่อ-สกุล	รหัสประจำตัว	ลายมือชื่อ
นายดิษพงษ์ ชชทรงธรรม	5030167521	.....
นายคมกรณ์ อัสวมณี	5031011421	.....
อาจารย์ที่ปรึกษาโครงการ		ลายมือชื่อ
ผศ.ดร.อรรถวิทย์ สุดแสง		.....
อาจารย์ที่ปรึกษาร่วมโครงการ		ลายมือชื่อ
อาจารย์ ดร.นัทที นิภานันท์		.....

## ที่มาและความสำคัญของปัญหา

ในปัจจุบันเทคโนโลยี Augmented reality กำลังเข้ามามีบทบาทมากขึ้นต่อการดำรงชีวิตของเรา ทั้งจากทาง Smart phone ที่เริ่มมีการนำเทคโนโลยีนี้เข้ามาแก้ไขปัญหาดังกล่าว เช่น การผสานข้อมูลสถานที่บนแผนที่และตำแหน่งจาก GPS เข้ากับภาพถ่ายจากกล้องวิดีโอเพื่อบอกทิศทาง [3][4], การรู้จำและให้ข้อมูลสถานที่สำคัญโดยการวิเคราะห์จากภาพวิดีโอและตำแหน่งจาก GPS [5][6] เป็นต้น นอกจากนี้ยังมีการใช้เทคโนโลยีนี้บนอุปกรณ์อื่นๆ เช่น การเติมลูกเล่นบนภาพถ่ายให้สอดคล้องกับใบหน้าคน [7], การเสริม effect บนเกมส์กระดานเพื่อให้เกมมีความสมจริงและสนุกยิ่งขึ้น [8] เป็นต้น

ยิ่งไปกว่านั้นยังมีการผสานวัตถุเสมือนเข้ากับโลกจริงด้วยการวิเคราะห์ภาพจากโลกจริงก่อนแล้วแสดงภาพวัตถุเสมือนซ้อนลงไปเหมือนกับว่ามีวัตถุจริงอยู่ตรงนั้นเช่น เกม Kweekie [9], Parallel Tracking and Mapping for Small AR Workspaces [10][11] เป็นต้น ซึ่งแม้ว่าจะมีการพยายามแสดงวัตถุเสมือนให้สมจริงด้วยการแสดงเป็นภาพสามมิติ แต่อย่างไรก็ดี เราก็ไม่สามารถใช้มือจับวัตถุที่แสดงนั้นได้ แม้จะมีผลงานหลายชิ้นที่พยายามทำส่วนติดต่อผู้ใช้ให้สามารถใช้งานได้ในการติดต่อกับระบบผ่านทางกล้องเว็บแคมเช่น Sixth Sense [12], Hand Gesture Recognition for interactive interfaces [13], HandVu [14] เป็นต้น ซึ่งจะเน้นการวิเคราะห์แยกแยะท่ามือเพื่อทำการควบคุมระบบ แต่ผลงานเหล่านี้ เน้นการควบคุมระบบในสองมิติ ทำให้ไม่สามารถควบคุมวัตถุที่เป็นสามมิติได้อย่างสมจริงและเป็นธรรมชาติ [15]

โครงการ “ระบบควบคุมวัตถุเสมือน (Augmented Reality Object Manipulation System)” นี้จึงถูกจัดทำขึ้น เพื่อทำการศึกษาระบบควบคุมวัตถุเสมือนด้วยมือในสามมิติผ่านทางกล้องเว็บแคมโดยเน้นการจำลองการปฏิสัมพันธ์ระหว่างมือและวัตถุเสมือนในเชิงกายภาพ คือทำให้เราสามารถจับและเคลื่อนย้ายวัตถุเสมือนในสามมิติได้ ทั้งนี้เพื่อเป็นการเสริมความสมจริงของวัตถุดังกล่าวอันจะทำให้เกิดประโยชน์กับงานด้าน Augmented reality โดยเฉพาะในด้านการนำเทคโนโลยีดังกล่าวมาใช้ในการสร้างระบบวัตถุเสมือนจริงที่ต้องการให้มีปฏิสัมพันธ์กับคนผ่านการจับด้วยมือ เช่น ระบบวัตถุเสมือนสำหรับทำการทดลองทางวิทยาศาสตร์, ระบบวัตถุเสมือนของเล่นตัวต่อเสริมทักษะ เป็นต้น โดยผลสำเร็จจะเป็นโปรแกรมระบบวัตถุเสมือนจริงที่สามารถปรับให้แสดงระบบวัตถุใดๆก็ได้ตามกับชุดโมเดลที่โหลดให้กับระบบ

## วัตถุประสงค์

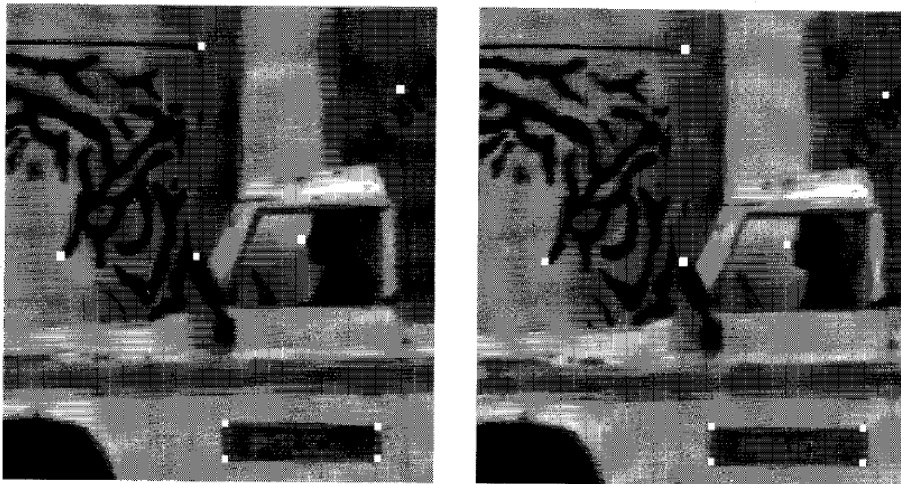
เพื่อให้มือสามารถ จับ, พลัก, บิด กับวัตถุเสมือนใน Augmented reality ได้

## ทฤษฎีที่เกี่ยวข้อง

### Stereopsis

กระบวนการ Stereopsis เป็นกระบวนการวิเคราะห์ข้อมูลบน โครงสร้างสามมิติและวิเคราะห์ระยะห่างระหว่างฉากถึงวัตถุ โดยใช้ภาพตั้งแต่สองภาพขึ้นไป ซึ่งได้มาจากมุมมองที่แตกต่างกัน โดยกระบวนการ Stereopsis นั้นเลียนแบบกระบวนการที่ใช้ในการทำงานของสมองที่รับภาพมาจากตาทั้งสองข้างแล้วประมวลผลให้กลายเป็นภาพสามมิติ กระบวนการ Stereopsis ต้องแก้ปัญหาสองอย่างคือ

1. ปัญหาการจับคู่กันของจุดในแต่ละภาพที่อ้างอิงไปยังจุดเดียวกันในฉาก (The Correspondence Problem) เช่น จุดในภาพจากตาด้านซ้าย ตรงกับจุดใดในภาพจากตาด้านขวา



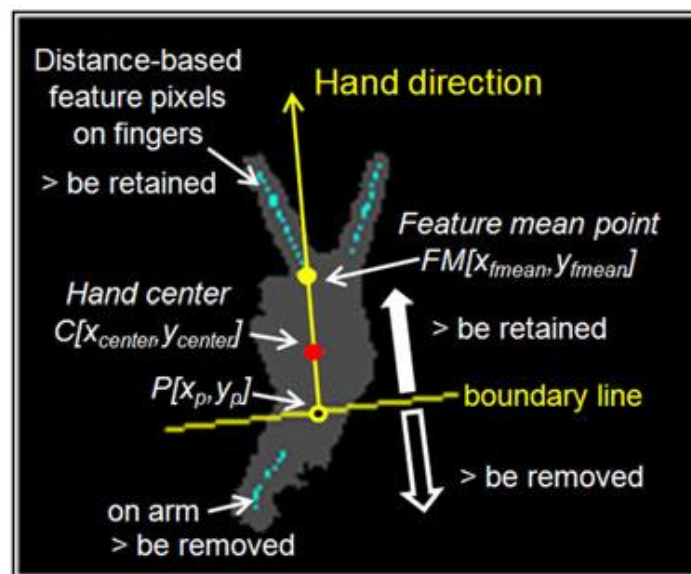
รูปแสดงปัญหาการจับคู่กันของจุดในแต่ละภาพที่อ้างอิงไปยังจุดเดียวกันในฉาก (Trucco and Verri, Introductory Techniques for 3-D Computer Vision)

2. ปัญหาการนำจุดทั้งหมดที่ได้จากการแก้ปัญหาในข้อที่ 1 มาระบุพิกัดสามมิติและสร้างเป็นโครงสร้างสามมิติ

## Hand feature extraction

การวิเคราะห์หาตำแหน่งของจุดสำคัญบนมือ ได้แก่ จุดกลางฝ่ามือ และปลายนิ้วทั้งหมด โดยอาศัยกระบวนการดังนี้

- Hand Segmentation คือการแยกส่วนของมือออกจากส่วนอื่นของรูป โดยใช้การตรวจจับสีผิว
- การหาจุดสำคัญด้วยเทคนิคตามงานวิจัย “Fast Hand Feature Extraction Based on Connected Component Labeling, Distance Transform and Hough Transform”(Le Dung and Makoto Mizukawa)



(Le Dung and Makoto Mizukawa, “Fast Hand Feature Extraction Based on Connected Component Labeling, Distance Transform and Hough Transform”)

## งานวิจัยที่เกี่ยวข้อง

- Real Time Hand Tracking and 3D Gesture Recognition for Interactive Interfaces using HMM [13]

งานวิจัยเกี่ยวกับการติดตามมือและการจำแนกท่าทางสามมิติแบบ Real Time สำหรับอินเตอร์เฟซแบบ Interactive โดยใช้ HMM (Hidden Markov Models) โดยจะตรวจจับและจำแนกท่าทางมือโดยให้ผู้ใช้สวมถุงมือสี ซึ่งจะทำได้ฝึกัดของมือผ่านทางสร้างโครงสร้างสามมิติ

- Fast Hand Feature Extraction Based on Connected Component Labeling, Distance Transform and Hough Transform [16]  
งานวิจัยเกี่ยวกับการสกัดข้อมูลของมือความเร็วสูงด้วยเทคนิค Connected Component Labeling, Distance Transform and Hough Transform ซึ่งใช้ภาพเพียงมุมมองเดียวและไม่ต้องมีตัวระบุตำแหน่งบนมือ (marker)

## ขอบเขตของโครงการ

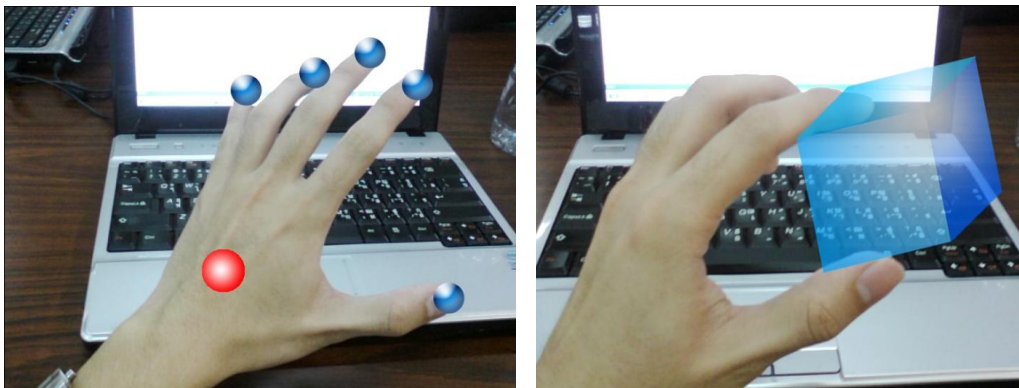
- เพื่อพัฒนาระบบวัตถุเสมือนจริงที่สามารถทำการปฏิสัมพันธ์กับมือในเชิงกายภาพในลักษณะของ Augmented reality (คือรับ input ผ่านกล้องและแสดงภาพวัตถุซ้อนลงบนภาพจากกล้อง)
- สร้างชุดโมเดลวัตถุเสมือนจริงตัวอย่างทดสอบ 1 ชุด คือชุดของเล่นตัวต่อเสริมทักษะ
- เนื่องจากใช้เว็บแคมในการพัฒนาจึงต้องการสภาวะแวดล้อมในการใช้งานที่มีแสงเพียงพอ
- Physic engine ที่ทำการพัฒนาขึ้นจะมีข้อจำกัดดังนี้
  - มีการคำนวณความสัมพันธ์เฉพาะระหว่างมือกับวัตถุ จะไม่พิจารณาการชนของกันของวัตถุเสมือน
  - วัตถุเสมือนเป็นรูปลูกบาศก์ (สามารถใช้วัตถุรูปทรงอื่นในลูกบาศก์ใส่ได้ในลักษณะของ bounding box)
 ทั้งนี้เพื่อให้งานมีประเด็นอยู่ในเรื่อง Computer vision เป็นหลัก

## รายละเอียดของการพัฒนา

การทำงานของระบบเริ่มจากการรับภาพจากกล้องเว็บแคมสองตัว (รูปที่ 1) ทำการคำนวณหามือและตำแหน่งในสามมิติของจุดสำคัญของมือซึ่งประกอบไปด้วยปลายนิ้วทุกนิ้ว, ศูนย์กลางฝ่ามือ เป็นต้น (รูปที่ 2) ระบบนำตำแหน่งจุดสำคัญบนมือที่ได้ไปป้อนให้กับ Physic engine ซึ่งจะทำการคำนวณความสัมพันธ์เชิงกายภาพระหว่างจุดสำคัญบนมือและวัตถุเสมือนสามมิติที่แสดงด้วย Computer graphic library เพื่อทำการเคลื่อนวัตถุเสมือนให้สอดคล้องกับการเคลื่อนไหวและตำแหน่งของมือ (รูปที่ 3)

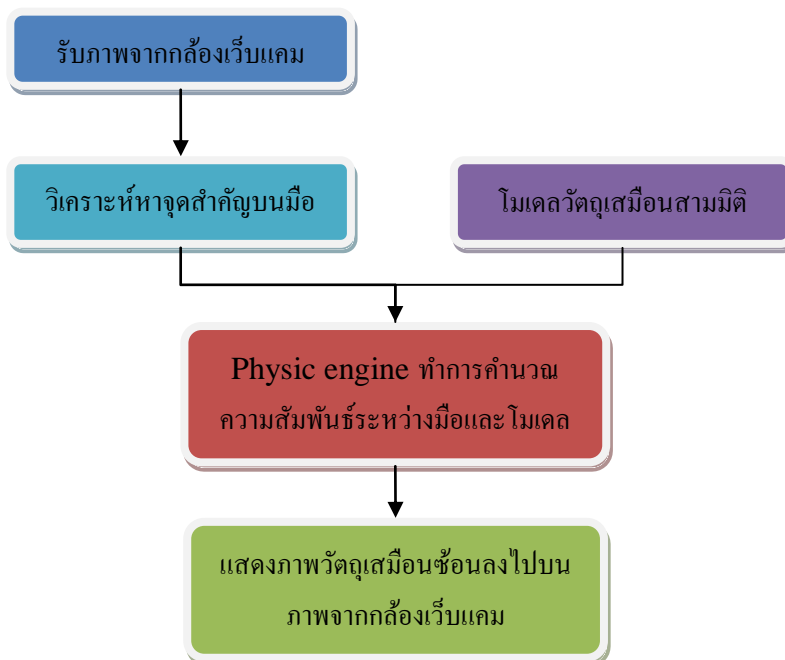


รูปที่ 1 แสดงภาพที่ได้มาจากกล้องสองตัวในต่างมุมมองกัน  
กล้องทางซ้าย(รูปซ้าย) และ กล้องทางขวา(รูปขวา)



รูปที่ 2 แสดงผลลัพธ์ของการหาจุดสำคัญของมือ      รูปที่ 3 แสดงตัวอย่างการจับวัตถุเสมือน

สรุปการทำงานของระบบได้ดังนี้



## เครื่องมือที่ใช้ในการพัฒนา

- Webcam
- OpenCV สำหรับการทำงานด้าน computer vision
- OpenGL สำหรับการแสดงวัตถุเสมือนสามมิติ
- Microsoft Visual Studio 2008 Express เป็นเครื่องมือในการพัฒนาระบบ

## รายละเอียดของโปรแกรมที่จะพัฒนา

### Input/output specification

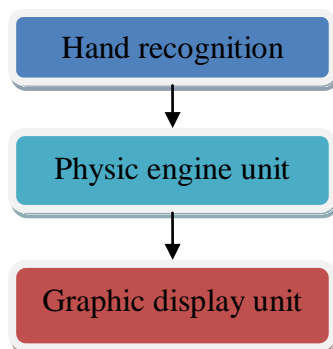
Input: ภาพวิดีโอจากกล้องเว็บแคมแสดงการเคลื่อนไหวของมือผู้ใช้

Output: การตอบสนองของระบบวัตถุเสมือนที่สอดคล้องกับมือ

### Functional Specification

- Hand recognition unit:  
หาพิกัดสามมิติของจุดสำคัญบนมือได้จากข้อมูลภาพของกล้องเว็บแคม  
เป้าหมาย: หาพิกัดได้ถูกต้องและต่อเนื่อง 80% แบบ real-time
- Physic engine unit:  
คำนวณปฏิสัมพันธ์เชิงกายภาพระหว่างมือและวัตถุเสมือนได้ใน 3 รูปแบบคือ จับ, ผลัก, บีค  
เป้าหมาย: ทำการจำลองความสัมพันธ์ทางกายภาพได้ใกล้เคียงความจริงแบบ real-time
- Graphic display unit:  
แสดงวัตถุซ้อนลงบนภาพจากเว็บแคมตามผลการคำนวณของ Physic engine ได้  
โดยจะแสดงวัตถุเสมือนในกรอบอ้างอิงของกล้องเสมอ (โครงการนี้จึงจะไม่มีการใช้ตัวอ้างอิงตำแหน่งในโลกจริง (marker))  
เป้าหมาย: แสดงภาพได้แบบ real-time

## โครงสร้างซอฟต์แวร์ (Design)



## ขั้นตอนการพัฒนาโปรแกรม

1. กำหนดขอบเขต เป้าหมาย และวัตถุประสงค์ของโครงการ
2. วางแผนการดำเนินงาน
3. แบ่งการดำเนินงานออกเป็น 3 ส่วนหลัก ๆ ดังนี้
  - ส่วนตรวจหามือ
  - ส่วน Physic Engine
  - ส่วนสำหรับแสดงผลสามมิติ
4. ศึกษาและพัฒนาโปรแกรมในแต่ละส่วน
5. รวมระบบย่อยทั้งสามเข้าด้วยกัน

ID	Task Name	Jul 2010	Aug 2010			Sep 2010			Oct 2010			Nov 2010			Dec 2010			Jan 2011									
		18/7	25/7	1/8	8/8	15/8	22/8	29/8	5/9	12/9	19/9	26/9	3/10	10/10	17/10	24/10	31/10	7/11	14/11	21/11	28/11	5/12	12/12	19/12	26/12	2/1	9/1
1	ศึกษาทฤษฎีและผลงานทางวิชาการที่เกี่ยวข้อง																										
2	ส่วน Hand tracking																										
3	ศึกษาและพัฒนาส่วน Hand feature extraction																										
4	ศึกษาและพัฒนาส่วน Stereopsis																										
5	ศึกษาพัฒนาการ Track มือแบบสามมิติ																										
6	ส่วน Physic Engine																										
7	ศึกษาการคำนวณ Collision detection																										
8	พัฒนา Physic engine																										
9	ส่วน 3D Display																										
10	ศึกษาการใช้งาน OpenGL																										
11	รวมทั้งสาม Module เข้าด้วยกัน																										
12	พัฒนาระบบวัตถุเสมือน																										
13	ปรับแต่งให้เกิดความเสมือนจริง																										

Gantt chart



## ประโยชน์ที่คาดว่าจะได้รับ

ในปัจจุบันมีการใช้เทคโนโลยี Augmented reality ในการเติมแต่งวัตถุเสมือนให้กับภาพจากกล้องเพื่อวัตถุประสงค์ต่างๆ แม้การแสดงผลจะสมจริงเพียงใด แต่เราก็ไม่สามารถทำการจับเคลื่อนวัตถุเหล่านั้นด้วยมือได้เหมือนกับวัตถุจริง ทำให้วัตถุเสมือนขาดความสมจริงและการใช้งานถูกจำกัดอยู่เพียงการมองเห็นและการโต้ตอบผ่านช่องทางอื่นๆที่ไม่ใช่การจับโดยตรง เช่น ผ่านทางเมาส์, จอสัมผัส เป็นต้น

โครงการนี้จึงจัดทำขึ้นเพื่อมุ่งเน้นที่จะแก้ปัญหาดังกล่าว ด้วยการทำระบบวัตถุเสมือนใน Augmented reality ที่สามารถจำลองการควบคุมเชิงกายภาพด้วยมือได้ เพื่อให้เราสามารถที่จะจับเคลื่อนย้ายวัตถุเหล่านั้นได้ด้วยมือเหมือนวัตถุอยู่ตรงนั้นจริงๆ ซึ่งจะทำให้วัตถุเสมือนมีความสมจริงมากยิ่งขึ้น และสามารถประยุกต์ใช้งานได้หลากหลายหลายมากยิ่งขึ้น

## เอกสารอ้างอิง

- [1] Emanuele Trucco and Alessandro Verri, “Introductory Techniques for 3-D Computer Vision”, Prentice Hall, 1998
- [2] Gary Bradski and Adrian Kaehler, “Learning OpenCV”, 1st edition, United State of America: O’Reilly, 2008
- [3] Wikitude, <http://www.wikitude.org/>
- [4] Austrian startup Mobilizy unveils augmented reality driving directions, <http://social.venturebeat.com/2009/08/28/austrian-startup-mobilizy-unveils-augmented-reality-driving-directions/>
- [5] Augmented Reality - Layar Reality Browser, <http://www.layar.com>
- [6] Layar, worlds first mobile Augmented Reality browser, [http://www.youtube.com/watch?v=b64\\_16K2e08](http://www.youtube.com/watch?v=b64_16K2e08)
- [7] New YouCam 3 Webcam Software, [http://www.cyberlink.com/products/youcam/overview\\_en\\_US.html](http://www.cyberlink.com/products/youcam/overview_en_US.html)
- [8] PIT Strategy, <http://www.augmented-reality-games.com/page.asp?pageid=215>
- [9] Kweekies, <http://www.intl3.net/kweekies-augmented-reality/en/>
- [10] Georg Klein and David Murray, “Parallel Tracking and Mapping for Small AR Workspaces”, Active Vision Laboratory, University of Oxford, 2007, <http://www.robots.ox.ac.uk/~gk/publications/KleinMurray2007ISMAR.pdf>
- [11] Parallel Tracking and Mapping (PTAM) news, <http://ewokrampage.wordpress.com/>
- [12] Pranav Mistry and Pattie Maes and Liyan Chang, “WUW - Wear Ur World - A Wearable Gestural Interface”, MIT Media Lab, [http://fluid.media.mit.edu:4000/assets/\\_pubs/wip177-mistry.pdf](http://fluid.media.mit.edu:4000/assets/_pubs/wip177-mistry.pdf)
- [13] C. Keskin, A. Erkan, L. Akarun, “REAL TIME HAND TRACKING AND 3D GESTURE RECOGNITION FOR INTERACTIVE INTERFACES USING HMM”, Boğaziçi University, 2009, <http://cs.nyu.edu/~naz/docs/icann.pdf>
- [14] Handvu Hand Gesture Recognition, <http://www.movesinstitute.org/~kolsch/HandVu/HandVu.html>
- [15] HandVu for virtual and augmented environments, 2004, [http://www.movesinstitute.org/~kolsch/HandVu/hv\\_ARVR.wmv](http://www.movesinstitute.org/~kolsch/HandVu/hv_ARVR.wmv)

- [16]Le Dung and Makoto Mizukawa, “Fast Hand Feature Extraction Based on Connected Component Labeling, Distance Transform and Hough Transform”, Shibaura Institute of Technology, 2009