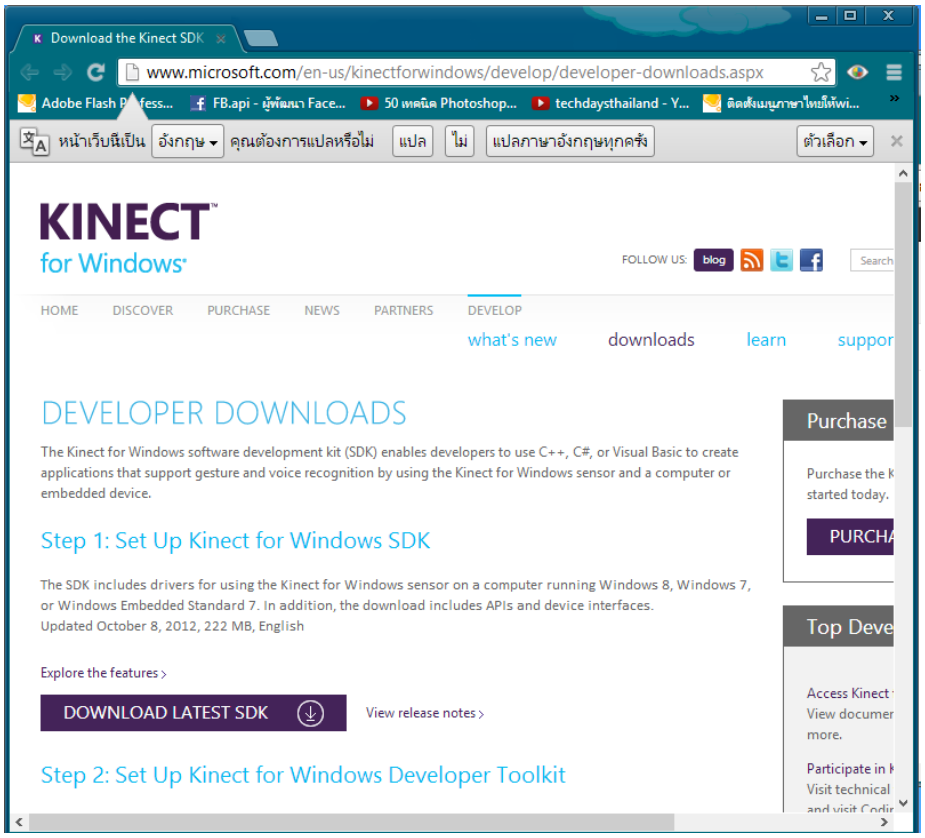
ภาคผนวก

ภาคผนวก ก
วิธีติดตั้ง Kinect for Windows SDK

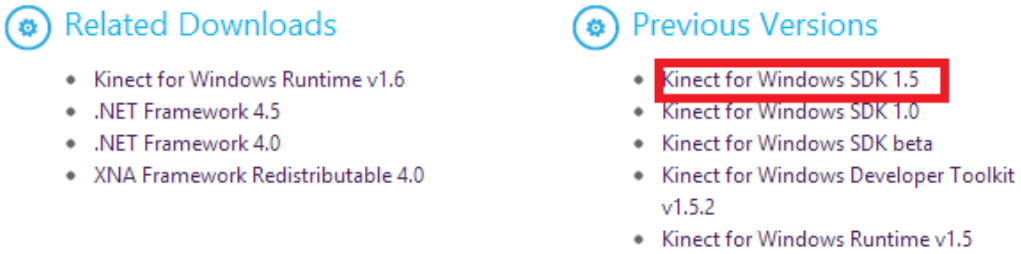
คู่มือการติดตั้ง Kinect for Windows sdk

1. ก่อนอื่นเข้าไป download ที่ลิงค์ http://www.microsoft.com/en-us/kinectforwindows/develop/developer-downloads.aspx



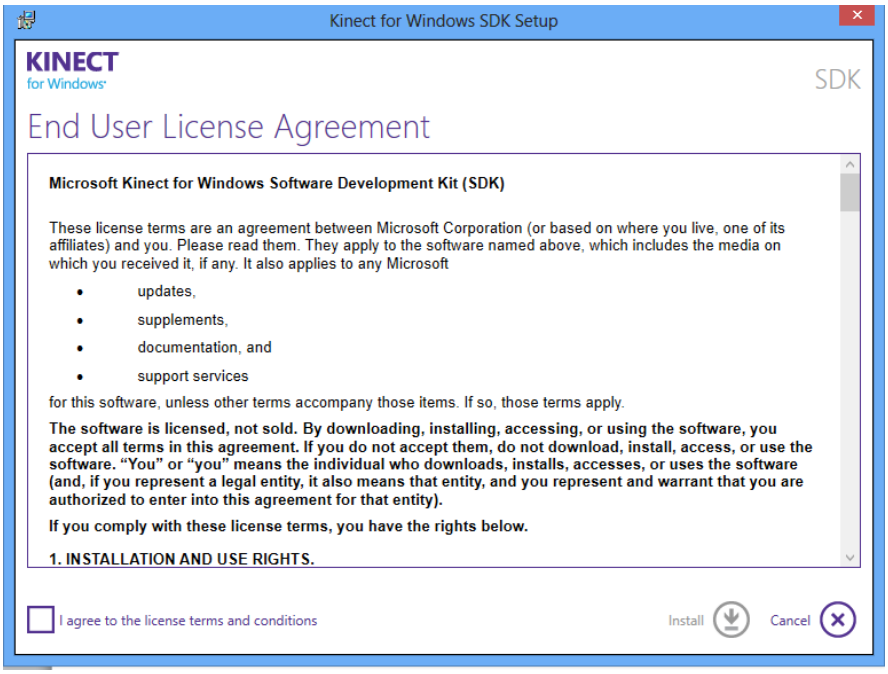
ภาพที่ 1 แสดงหน้าเว็ป Download

เลื่อนหา ข้อความดังภาพ แล้วคลิกเพื่อดาวน์โหลด แล้วจะได้ไฟล์ติดตั้ง .exe มา



ภาพที่ 2 แสดงหน้า URL Download

2. เปิดไฟล์ kinectSDK-1.5.exe ที่ดาวน์โหลดมาไว้บนเครื่อง ก็จะได้หน้าต่างดังภาพ

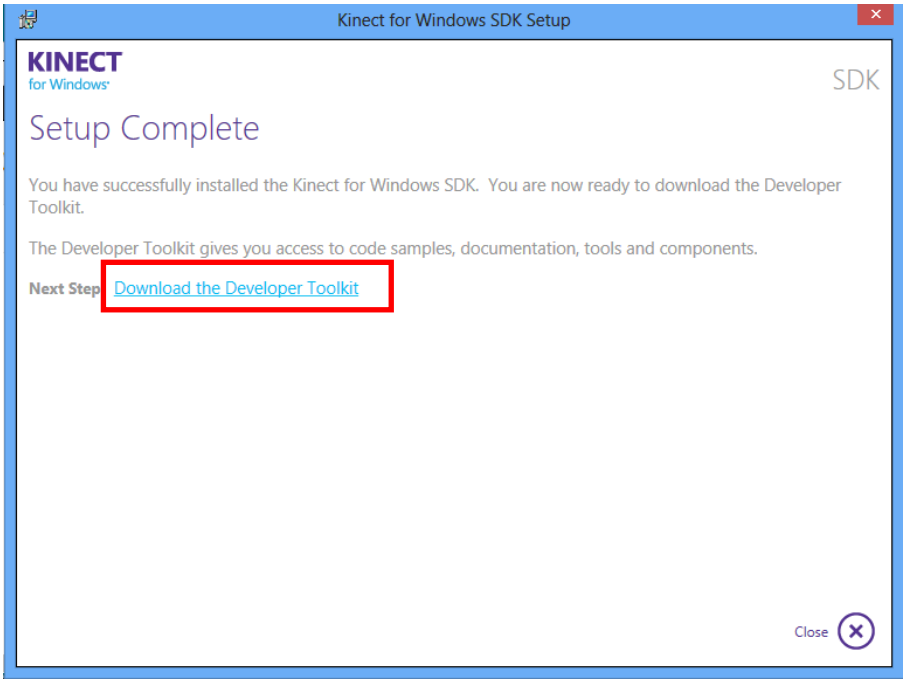


ภาพที่ 3 แสดงขั้นตอนการติดตั้ง

- เลือกเครื่องหมายถูก ในช่อง ตามภาพ
- จากนั้น คลิ๊ก ตรง Install ได้เลย เพื่อเริ่มทำการติดตั้งโปรแกรม

ภาพที่ 4 แสดงขั้นตอนการติดตั้ง

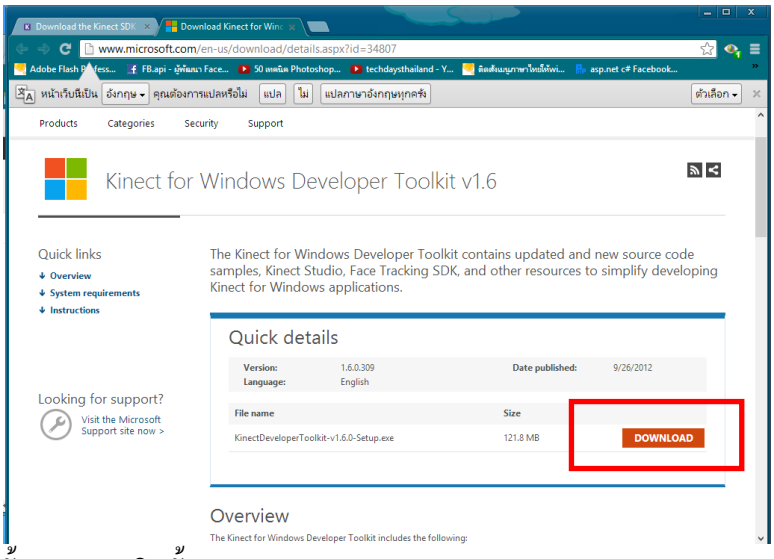
3. เมื่อเลือกที่ Install แล้ว ก็จะแสดงหน้าต่างดังภาพ



ภาพที่ 5 แสดงขั้นตอนการติดตั้ง

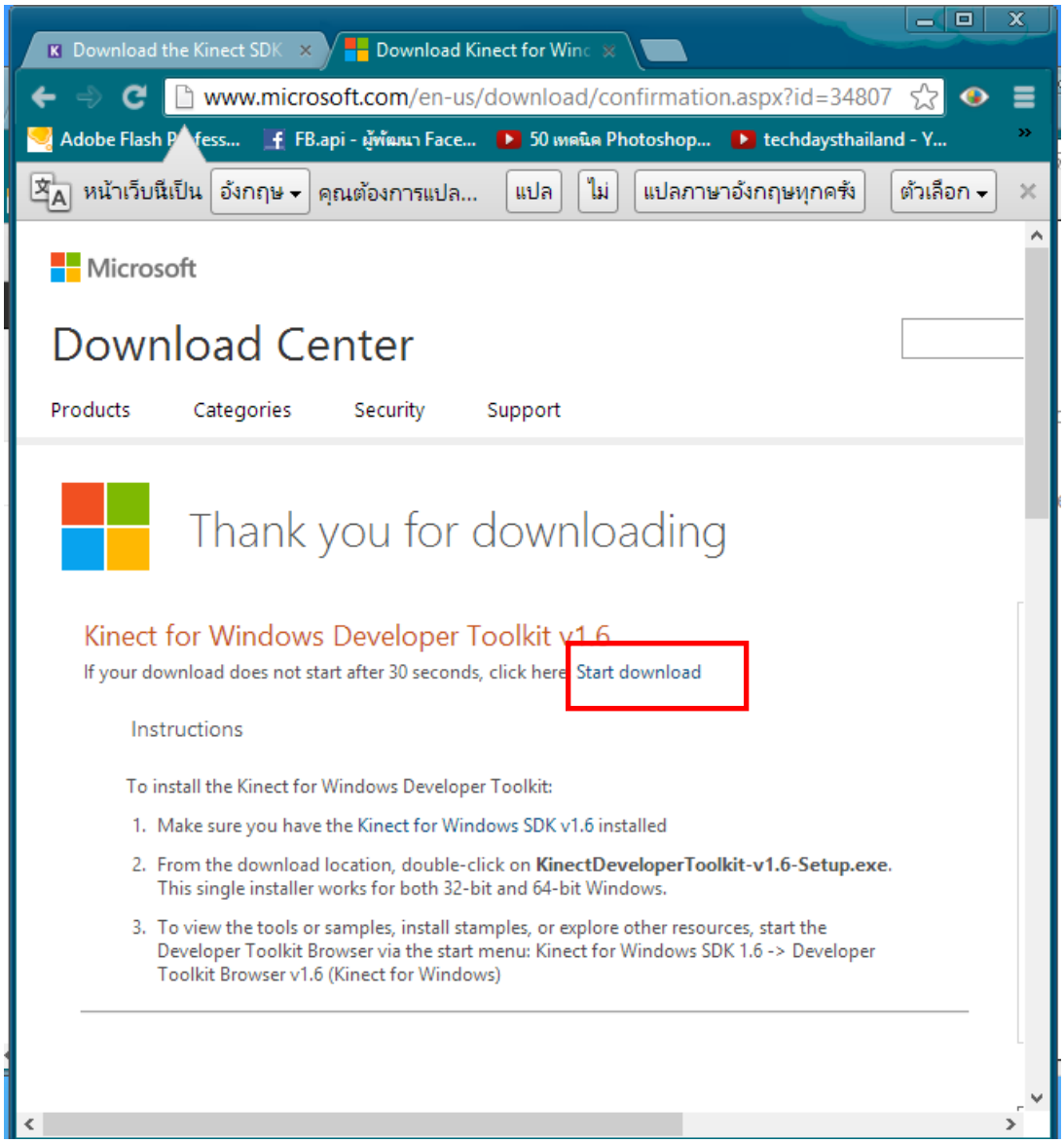
- เมื่อได้หน้าต่างตามนี้แล้วให้ คลิก Download the Developer Toolkit ตามภาพ เพื่อที่จะทำการ Download Kinect for Windows Developer Toolkit

4 เมื่อทำตามข้อ 3 แล้ว ก็จะมีหน้าต่างบน web browser ขึ้นมาเพื่อจะทำการ Download ต่อไป



ภาพที่ 6 แสดงขั้นตอนการติดตั้ง

5 คลิกที่ Start download เพื่อทำการ Download เป็นขั้นตอนสุดท้ายของการ Download Kinect for Windows Developer Toolkit



ภาพที่ 7 แสดงขั้นตอนการติดตั้ง

download เสร็จแล้วก็จะได้ file ติดตั้ง .exe

6 เมื่อ Download file (.exe) มาเก็บไว้ ที่เครื่องแล้วก็ทำการเปิดไฟล์นั้นขึ้นมาเลย เพื่อที่จะทำการติดตั้ง Kinect for Windows Developer Toolkit เมื่อเปิดขึ้นมาแล้วก็ จะได้เห็นหน้าต่างในภาพ

ภาพที่ 8 แสดงขั้นตอนการติดตั้ง

- เลือก เครื่องหมายถูก ในช่อง ตามภาพ
- คลิกที่ Install เพื่อเริ่มทำการติดตั้ง

ภาคผนวก ข
ตารางบันทึกผลการทดลองครั้งที่ 1

การทดลองครั้งที่ 1 จำนวน 30 คน ไม่มีเสียงรบกวน

ตารางที่ 1 ตารางบันทึกผลการทดลองกลุ่มที่ 1 Voice Only (ไม่มีเสียงรบกวน)

ลำดับที่	ชุดทดสอบที่	ระยะเวลาที่ใช้ทดลอง(s)	จำนวนครั้งที่ผิดพลาด	รายละเอียดของข้อผิดพลาด
1	C	21.44	2	ออกเสียง CH 11, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
2	B	19	2	ออกเสียง Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
3	C	17.9	0	
4	C	14.7	1	ออกเสียง CH 9 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
5	A	20.3	0	-
6	B	12.05	0	-
7	A	11.25	0	-
8	B	8.7	0	-
9	A	57.2	10	ออกเสียง CH 7, CH 3, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
10	B	11.8	0	-
11	C	18.7	3	ออกเสียง CH 9,CH 11 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
12	B	9.8	0	-
13	C	13.6	1	ออกเสียง CH 9 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
14	A	16.8	3	ออกเสียง CH 3, CH 7 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
15	B	11.19	0	-

16	A	15.4	1	ออกเสียง Ch 7 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
17	C	17.57	1	ออกเสียง CH 9 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
18	A	22.85	2	ออกเสียง CH 3 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
19	C	29.03	4	ออกเสียง Ch9,Volum Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
20	C	13.29	0	-
21	A	16.87	1	ออกเสียง Volum Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
22	B	21.16	1	ออกเสียง CH 11 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
23	C	41.03	4	ออกเสียง CH 11, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
24	A	16.76	0	-
25	B	22.49	0	-
26	C	25.39	2	ออกเสียง Volume Up, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
27	A	13.55	2	ออกเสียง Ch7 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
28	C	29.16	3	ออกเสียง CH 11, Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
29	A	11.86	0	-
30	A	18.78	0	-

ตารางที่ 2 ตารางบันทึกการทดลองกลุ่มที่ 2 Motion Detection Only (ไม่มีเสียงรบกวน)

ลำดับที่	ชุดทดสอบที่	ระยะเวลาที่ใช้ทดลอง(s)	จำนวนครั้งที่ผิดพลาด	รายละเอียดของข้อผิดพลาด
1	A	20.15	0	-
2	A	23.70	0	-
3	B	18.11	0	-
4	A	13.78	0	-
5	B	20.11	0	-
6	A	20.72	0	-
7	B	25.10	1	เลื่อนไป Detection Volume Up แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
8	A	22.89	0	-
9	C	19.14	0	-
10	A	19.50	0	-
11	B	34.41	2	เลื่อนไป Detection Volume Down แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
12	A	30.57	1	เลื่อนไป Detection CH 3 แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
13	B	20.02	0	-
14	C	38.14	0	-
15	A	20.02	0	-
16	B	25	0	-
17	A	22.44	0	-
18	B	33.43	1	เลื่อนไป Detection CH 7แล้วแต่ตัว Pointer อยู่ไม่

				ครบเวลา
19	A	12.85	0	-
20	A	15.50	0	-
21	B	18.98	0	-
22	A	21.94	0	-
23	A	15.77	0	-
24	B	19.92	0	-
25	B	25.94	1	เลื่อนไป Detection CH 7แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
26	B	29.34	1	เลื่อนไป Detection CH 7แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
27	B	14.81	0	-
28	B	15.38	0	-
29	C	22.19	1	เลื่อนไป Detection CH 9แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
30	B	20.29	0	-

ตารางที่ 3 ตารางบันทึกการทดลองกลุ่มที่ 3 A Combination (ไม่มีเสียงรบกวน)

ลำดับที่	ชุดทดสอบที่	ระยะเวลาที่ใช้ทดลอง(s)	จำนวนครั้งที่ผิดพลาด	รายละเอียดของข้อผิดพลาด
1	B	25.8	0	-
2	C	32.6	0	-
3	A	30.2	0	-
4	B	24.1	0	-
5	C	30.2	0	-
6	C	27.3	0	-
7	C	29.2	0	-
8	C	22.3	0	-
9	B	35.6	0	-
10	C	20.44	0	-
11	A	20.7	0	-
12	C	35.5	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
13	A	26.7	0	-
14	B	18.9	0	-
15	C	20	0	-
16	C	23.7	0	-
17	B	19.23	0	-
18	C	21.17	0	-
19	B	15.08	0	-

20	B	22.17	0	
21	C	17.68	0	-
22	B	25.23	0	-
23	B	29.26	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
24	C	27.68	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
25	C	25.01	0	-
26	A	16.36	0	-
27	C	15.24	0	-
28	A	19.16	0	-
29	B	21.63	0	-
30	C	20.25	0	-

<u>การทดลองครั้งที่ 1 จำนวน 30 คน มีเสียงรบกวน</u>

ตารางที่ 4 ตารางบันทึกผลการทดลองกลุ่มที่ 1 Voice Only (มีเสียงรบกวน)

ลำดับที่	ชุดทดสอบที่	ระยะเวลาที่ใช้ทดลอง(s)	จำนวนครั้งที่ผิดพลาด	รายละเอียดของข้อผิดพลาด
1	A	30.35	3	ออกเสียง Ch3, Volume Down, Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
2	C	20.14	2	ออกเสียง Ch9 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
3	B	19.53	1	ออกเสียง Ch5 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
4	C	47.14	5	ออกเสียง Ch9, Ch11, Volume Up, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
5	B	17.56	1	ออกเสียง Volume Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
6	C	18.55	1	ออกเสียง Ch 11 แล้วไม่มีการตอบสนองหรือตอบสนอง
7	B	18.53	1	ออกเสียง Ch 5 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
8	C	32.89	2	ออกเสียง Volume Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
9	A	24.90	2	ออกเสียง Ch7 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
10	B	19.45	1	ออกเสียง Ch 9 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
11	A	36.63	5	ออกเสียง Ch3, Ch7, Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
12	C	26.61	3	ออกเสียง Ch11, Volune Donw แล้วไม่มีการตอบสนองหรือตอบสนองผิด

13	B	21.61	3	ออกเสียง Ch 5, Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
14	C	19.66	1	ออกเสียง Ch9 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
15	A	26.63	2	ออกเสียง Ch3 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
16	B	21.91	1	ออกเสียง Volume Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
17	C	52.38	7	ออกเสียง Ch9, Volume Donw,Ch11,Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
18	A	36.73	6	ออกเสียง Ch3, Ch7, Volume Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
19	B	26.80	3	ออกเสียง Ch5,Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
20	A	42.20	8	ออกเสียง Ch3, Ch7, Volume Down, Volume Up, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
21	C	22.56	1	ออกเสียง Volume Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
22	B	19.58	1	ออกเสียง Ch9, Volume Down Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
23	C	32.06	4	ออกเสียง Ch9, Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
24	A	21.72	2	ออกเสียง Ch7 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
25	C	71.05	11	ออกเสียง Ch9, Ch11, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด

26	B	16.69	1	ออกเสียง Volume Up Up แล้วไม่มีการตอบสนอง หรือตอบสนองผิด
27	A	62.25	9	ออกเสียง Ch3, Ch7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
28	C	19.81	1	ออกเสียง Ch11 แล้วไม่มีการตอบสนองหรือ ตอบสนองผิด
29	B	20.07	1	ออกเสียง Ch5 แล้วไม่มีการตอบสนองหรือ ตอบสนองผิด
30	A	37.81	4	ออกเสียง Ch7 Volume Up, Volume Down

ตารางที่ 5 ตารางบันทึกผลการทดลองกลุ่มที่ 3 A Combination (มีเสียงรบกวน)

ลำดับที่	ชุดทดสอบที่	ระยะเวลาที่ใช้ทดลอง(s)	จำนวนครั้งที่ผิดพลาด	รายละเอียดของข้อผิดพลาด
1	B	38.73	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
2	A	56.21	8	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
3	C	40.20	0	-
4	A	38.56	3	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
5	C	26.45	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
6	B	38.13	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
7	C	35.69	3	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
8	A	28.29	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
9	C	31.82	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
10	B	40.71	3	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
11	A	58.45	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
12	B	41.75	4	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
13	C	28.19	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
14	B	41.84	3	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
15	A	23.68	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง

16	C	21.89	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
17	A	29.60	3	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
18	B	33.99	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
19	C	27.75	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
20	A	33.36	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
21	C	24.23	0	-
22	B	26.54	0	-
23	A	45.43	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
24	B	40.81	6	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
25	A	22.26	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
26	C	25.26	0	-
27	B	32.17	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
28	C	27.94	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
29	A	28.39	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
30	B	34.51	4	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด

ภาคผนวก ค
ตารางบันทึกผลการทดลองครั้งที่ 2

การทดลองครั้งที่ 2 จำนวน 30 คน มีเสียงรบกวน

ตารางที่ 6 ตารางบันทึกผลการทดลองครั้งที่ 2 มีเสียงรบกวน (70 dB)

ลำดับที่	ลำดับการควบคุม	เทคนิคที่ใช้ควบคุม	ระยะเวลาที่ใช้ทดลอง (นาที)	จำนวนครั้งที่ผิดพลาด	รายละเอียดของข้อผิดพลาด
1	D	Voice	1.27	10	ออกเสียง Ch3, Ch7 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.44	2	เลื่อนไป Ch3,Ch5 แล้วแต่ตัว Pointer อยู่ไม่ครบเวลา
		Combination	0.50	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
2	C	Voice	0.56	10	ออกเสียง Ch3, Vol up, Ch 7, TV Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.44	-	-
		Combination	1.09	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
3	B	Combination	0.47	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
		Motion	0.38	-	-
		Voice	0.47	5	ออกเสียง TV ON, Ch5, Vol Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
4	A	Motion	1.03	-	-
		Combination	1.41	7	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Voice	1.34	15	ออกเสียง TV ON, Vol Up, Ch3, Tv On, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
5	E	Voice	1.53	11	ออกเสียง TV ON, Ch3, Ch5 แล้วไม่มี

					การตอบสนองหรือตอบสนองผิด
		Combination	1.57	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.50	-	-
6	F	Motion	0.49	-	-
		Voice	1.07	6	ออกเสียง Ch3, Vol Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	2.11	7	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
7	B	Combination	1.36	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	2.05	-	-
		Voice	1.33	7	ออกเสียง Vol Up, Ch3, Vol Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
8	E	Voice	0.58	5	ออกเสียง Ch3, Ch9, TVOn แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	1.18	7	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	1.10	-	-
9	D	Combination	0.53	6	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Voice	1.23	10	ออกเสียง Ch5, Ch9, Ch3, Ch 11, Volume Up, TV On, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.35	-	-

10	F	Motion	0.30	-	-
		Voice	0.57	3	ออกเสียง Volume Down, Volume Up แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	1.05	4	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
11	E	Voice	2.13	10	ออกเสียง Tv On, Ch 5, Ch9, Ch7 Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	1.00	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
		Motion	0.40	-	-
12	E	Voice	1.23	11	ออกเสียง Tv On, Ch 5, Ch9, Ch7, Ch3 Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	1.07	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
		Motion	0.32	-	-
13	D	Combination	1.09	2	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Voice	0.49	2	ออกเสียง Ch3 แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.36	-	-
14	D	Combination	1.36	4	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Voice	1.50	7	ออกเสียง Tv On, Ch5, Ch9, Ch7, Volume Down แล้วไม่มีการ

					ตอบสนองหรือตอบสนองผิด
		Motion	0.32	-	-
15	B	Combination	1.20	5	Ch5, Ch9, Ch7, Volume Up
		Motion	0.40	-	-
		Voice	3.15	10	ออกเสียง Tv On, Ch5, Ch3, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
16	E	Voice	0.58	2	ออกเสียง Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	0.45	2	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.31	-	-
17	F	Motion	0.35	-	-
		Voice	2.11	8	ออกเสียง Tv On, Ch3, Ch7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	0.48	-	-
18	C	Voice	1.19	4	ออกเสียง Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.56	-	-
		Combination	1.24	-	
19	F	Motion	0.54	-	-
		Voice	3.38	15	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	1.23	-	-
20	A	Motion	0.51	-	-
		Combination	1.04	-	-

		Voice	1.11	9	ออกเสียง Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
21	D	Combination	1.32	3	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Voice	3.43	15	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.49	-	-
22	A	Motion	0.49	-	-
		Combination	1.10	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
		Voice	2.15	8	ออกเสียง Ch9, Ch3, Ch 7, Volume Up, Volume Down แล้วไม่มีการตอบสนองหรือตอบสนองผิด
23	F	Motion	0.49	-	-
		Voice	2.55	12	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Combination	1.28	4	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
24	A	Motion	0.57	-	-
		Combination	0.59	5	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Voice	4.00	16	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv

					Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
25	C	Voice	3.15	13	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.49	-	-
		Combination	0.55	-	-
26	B	Combination	0.53	3	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.45	-	-
		Voice	4.00	16	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
27	A	Motion	0.50	-	-
		Combination	0.59	2	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
		Voice	0.52	-	-
28	C	Voice	2.45	9	ออกเสียง Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.45	-	-
		Combination	1.15	1	ออกเสียง OK แล้ว Pointer หลุดการแถบเปลี่ยนช่อง
29	B	Combination	2.17	4	Tv On, Volume Down, Tv Off
		Motion	0.58	-	-
		Voice	3.19	16	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch

					7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
30	C	Voice	3.05	12	ออกเสียง Tv On, Ch5, Ch9, Ch3, Ch 7, Volume Up, Volume Down, Tv Off แล้วไม่มีการตอบสนองหรือตอบสนองผิด
		Motion	0.54	-	-
		Combination	1.26	3	ออกเสียง OK แล้วไม่มีการตอบสนองหรือตอบสนองผิด

ภาคผนวก ง

ผลงานวิจัยในการประชุมวิชาการ การบริหารและการจัดการครั้งที่ 8 “การบริหารจัดการแห่งโลกพลวัต” โดยมหาวิทยาลัยธุรกิจบัณฑิตย์ ร่วมกับเครือข่ายวิจัยประชาชื่น

เทคนิคสำหรับการควบคุมอินเทอร์เน็ตทีวีโดยคีเนค ด้วยการผสมผสานระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว

A Technique for Internet Television Control using Kinect Combination of Voice Recognition and Motion Detection Approach

ภูวดล ศิริกองธรรม[1] และ วรสิทธิ์ ชูชัยวัฒนา[2]

บทคัดย่อ

การรู้จำเสียงหรือการตรวจสอบการเคลื่อนไหว ได้ถูกนำเข้ามาช่วยทำให้การควบคุมอินเทอร์เน็ตทีวี มีประสิทธิภาพและสะดวกสบายมากขึ้น แต่เทคโนโลยีดังกล่าวที่กล่าวมานั้น ต่างมีจุดเด่นและข้อด้อยในตัว ผู้วิจัยจึงเสนอเทคนิคที่ใช้ควบคุมอินเทอร์เน็ตทีวีแบบใหม่ด้วยการผสมผสานจุดเด่นระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว เพื่อทำให้การควบคุมอินเทอร์เน็ตทีวีมีประสิทธิภาพมากยิ่งขึ้น ในการทำการทดลองเพื่อวัดประสิทธิภาพ ผู้วิจัยทำการเปรียบเทียบเทคนิคที่นำเสนอ ซึ่งได้แก่การผสมผสานระหว่างการรู้จำเสียงและท่าทาง ในสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวน ซึ่งค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 เดซิเบล และสภาพแวดล้อมที่มีเสียงรบกวน โดยกำหนดให้เสียงรบกวนอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยิน รับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่ เฉลี่ย 70 เดซิเบล จากการทดลอง พบว่าขั้นตอนและวิธีการควบคุมอินเทอร์เน็ตทีวีด้วยการผสมผสานระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว สามารถทำงานได้อย่างถูกต้องทั้งในสภาพแวดล้อมทั้งสองแบบ

คำสำคัญ-- การรู้จำเสียง; ตรวจสอบการเคลื่อนไหว; อินเทอร์เน็ตทีวี;

Abstract

Voice recognition techniques and motion detection techniques were applied to improve an effectiveness of an internet television controlling task and make the task more convenient.

---

[1]นักศึกษาระดับบัณฑิตศึกษาสาขาวิศวกรรมเว็บ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยธุรกิจบัณฑิตย์

[2]อาจารย์ประจำสาขาวิชาวิศวกรรมเว็บ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยธุรกิจบัณฑิตย์

These techniques, however, provide some benefit and drawback to the task. A combination of voice recognition and motion detection techniques was proposed to further improve an effectiveness of the controlling task. To evaluate the proposed technique, a series of experiments were conducted. The proposed technique was compared with a voice recognition technique only and motion detection techniques only. The experiments were conducted in two different environments, which are a room with an average of 46 dB sound level and a room with an average of 70 dB sound level. The result of experiment showed that the combination of voice recognition and motion detection technique provide a better performance in both environments.

Keyword-- Speech recognition; Motion Detection; Internet TV;

1. บทนำ

1.1 ที่มาและความสำคัญของปัญหา

เทคโนโลยีเป็นส่วนสำคัญอย่างมากที่จะทำให้การดำรงชีวิตอยู่ของมนุษย์ สะดวกสบายมากขึ้น โทรทัศน์ก็เป็นอีกหนึ่งของเทคโนโลยีที่ช่วยเพิ่มความสามารถในการรับรู้ข่าวสาร และ ให้ความบันเทิง กับผู้ใช้ตลอดระยะเวลาที่ผ่านมาโทรทัศน์ได้ถูกพัฒนาอย่างต่อเนื่องในเรื่องของการแสดงผลภาพ แต่ก็คงยังใช้วิธีการควบคุมโดย เครื่องควบคุมระยะไกล (Remote Control) ในการควบคุมโทรทัศน์มาอย่างยาวนาน ซึ่งเป็นปัญหาในเรื่องของการควบคุมตัวโทรทัศน์ เนื่องจากรีโมทที่ใช้ควบคุมโทรทัศน์เป็นอีกหนึ่งอุปกรณ์ที่แยกจากตัวโทรทัศน์เองและจำเป็นต้องใช้แหล่งพลังงานเสริม และตัวรีโมทที่แยกออกมานั่นอาจมีการสูญหายทำให้การควบคุมโทรทัศน์เสียไป

อย่างไรก็ตามในปัจจุบันเริ่มมีการนำเทคโนโลยีการรู้จำเสียงและท่าทางมาใช้แทนที่สำหรับเครื่องควบคุมระยะไกล (Remote Control) แต่ก็ยังไม่เป็นที่นิยมเท่าที่ควรทั้งในเรื่องของฮาร์ดแวร์ที่มีราคาสูงและ เทคโนโลยีทั้งสองมีข้อด้อยกล่าวคือ เทคโนโลยีการรู้จำเสียงนั้นหากมีเสียงเข้ามารบกวนจะทำให้การทำงานเกิดความผิดพลาด เทคโนโลยีตรวจสอบการเคลื่อนไหวนั้นต้องใช้เวลาค่อนข้างมากในการสั่งงาน แต่ในปัจจุบันเทคโนโลยีทางตัวรับรู้ภาพและเสียงได้ถูกพัฒนาไปอีกระดับหนึ่ง โดยทางบริษัท ไมโครซอฟท์สามารถพัฒนาอุปกรณ์คีเนค (Kinect)ซึ่งเป็นอุปกรณ์ที่สามารถรับรู้ข้อมูลมิติความลึกของภาพและเสียงได้ในเวลาเดียวกัน ทำให้เพิ่มศักยภาพในการใช้ควบคุมโทรทัศน์ด้วยการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว

ดังนั้นงานวิจัยนี้ จึงเสนอเทคนิคในการควบคุมอินเทอร์เน็ตทีวี โดยใช้การผสมผสานระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว ผ่านอุปกรณ์คีเนค เพื่อทดสอบถึงความ

ถูกต้องและความเร็วในการควบคุมอินเทอร์เน็ตทีวีภายใต้สภาพแวดล้อมที่มีเสียงรบกวนและไม่รบกวน

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาการใช้งานกล้อง Kinect ในการตรวจจับการเคลื่อนไหว และการรู้จำเสียงของมนุษย์

2. เพื่อเสนอเทคนิควิธีที่ใช้ควบคุมอินเทอร์เน็ตทีวีแบบใหม่ด้วยการผสมผสานระหว่างการรู้จำเสียงและการตรวจจับการเคลื่อนไหว

3. เพื่อทดสอบประสิทธิภาพของวิธีการควบคุมอินเตอร์เน็ตทีวี ทั้งสามวิธี คือการรู้จำเสียง การตรวจจับการเคลื่อนไหว และการผสมผสานการรู้จำเสียงและการเคลื่อนไหว

4. เพื่อเสนอโปรแกรมแสดงผลการควบคุมอินเตอร์เน็ตทีวี

1.3 ขอบเขตของการศึกษา

วิทยานิพนธ์นี้ประกอบไปด้วย สองส่วนหลักๆ คือการออกแบบและกระประเมินผล ในส่วนแรกคือการออกแบบเทคนิคสำหรับการควบคุมอินเทอร์เน็ตทีวีโดย คีเนคด้วยการผสมผสานระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหวมาใช้กับระบบอินเตอร์ทีวีที่จำลองขึ้น ในส่วนที่สองเป็นการทดลองประเมินผลระบบด้วยเทคนิคที่นำเสนอและออกแบบไว้ โดยกลุ่มทดสอบจำนวน 60 คนนักศึกษาชั้นปี 1 – 3 ช่วงอายุประมาณ 19 – 21 ปี โดยจะแบ่งการทดสอบเป็น 2 กลุ่ม กลุ่มล่ะ 30 คนโดยให้กลุ่มแรกทำการทดสอบในสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวนค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB และกลุ่มที่สองอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเตอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB แล้วประเมินผลในเรื่องของความเร็วและความถูกต้องในการทำงาน

1.4 ความรู้เบื้องต้นและงานวิจัยที่เกี่ยวข้อง

1.4.1 คีเนค (Kinect) เป็นอุปกรณ์รับรู้การเคลื่อนไหวที่ใช้การทำงานผสมผสานกันระหว่างฮาร์ดแวร์และซอฟท์แวร์ ฟังก์ชันหลักของ คีเนค[1] มีสองฟังก์ชันคือ สร้างภาพเคลื่อนไหวสามมิติของวัตถุในมุมมองที่กำหนด และแยกแยะมนุษย์ออกจากวัตถุเหล่านั้นได้ ในปัจจุบัน คีเนคมีความละเอียดของภาพอยู่ที่ 640x480 pixel และสามารถทำงานได้ที่ 30 frames ต่อวินาที สำหรับ Hardware ของ คีเนคสามารถแบ่งออกเป็น ส่วนประกอบสำคัญ ได้แก่

ภาพที่ 1 สถาปัตยกรรมคีเนค

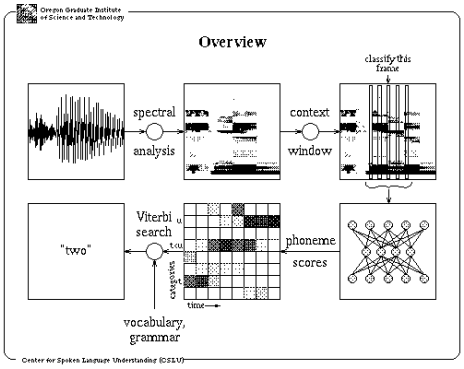
Color VGA video camera เป็นกล้องถ่ายภาพเคลื่อนไหวซึ่งช่วยในการจดจำใบหน้า (Face Recognition) และการตรวจจับลักษณะเด่นอื่นๆโดยใช้ความสามารถในการตรวจจับองค์ประกอบของสีทั้งสามสีอันได้แก่ สีแดง สีเขียว และสีน้ำเงิน

Depth sensor ประกอบด้วยการทำงานร่วมกันของต้นกำเนิดแสงอินฟราเรด (Infrared Projector) และตัวรับรู้แบบ monochrome CMOS (Complimentary Metal-Oxide Semiconductor) ซึ่งทำหน้าที่รับแสงอินฟราเรดที่ถูกสะท้อนกลับมาจากวัตถุ จากนั้นทำการวัดเวลาในการเดินทาง (Time of Flight) แสงอินฟราเรดนี้ใช้หลักการทำงานเช่นเดียวกับโซนาร์คือ หากรู้ระยะเวลาที่แสงอินฟราเรดใช้ในการเดินทางไปกลับก็จะสามารถคำนวณระยะห่างระหว่างตัวรับรู้ความลึกกับวัตถุได้ โดยการทำงานด้วยความเร็วแสงหลายๆรอบทำให้สามารถระบุระยะห่างที่แน่นอนได้โดยไม่ต้องคำนึงถึงสภาพแสงสว่าง

Multi-array microphone เป็นอาเรย์ของไมโครโฟน 4 ตัวซึ่งสามารถแยกแยะเสียงของผู้ใช้ออกจากเสียงรบกวนภายในห้องได้

1.4.2 ระบบการรู้จำเสียง (Speech recognition) [2] หมายถึง การทำให้คอมพิวเตอร์สามารถที่จะฟังคำพูดและตัดสินใจได้ว่าคำพูดนั้นเป็นคำว่าอะไรหรืออีกความหมายหนึ่งคือการนำ File Audio ที่บรรจุเสียงพูดนำมาแปลงเป็น Text ได้

Automatic Speech Recognition (ASR) เป็นเทคโนโลยีที่เกิดขึ้นเพื่อที่จะให้คอมพิวเตอร์สามารถแยกแยะคำพูดต่างๆที่มนุษย์สามารถพูดใส่อุปกรณ์ไมโครโฟนหรือเครื่องโทรศัพท์หรืออื่นๆเท่าที่จะเป็นไปได้ “ Holy grail ” of ASR Search เป็นโปรแกรมที่อนุญาตให้คอมพิวเตอร์เข้าใจคำศัพท์ทุกคำอย่างถูกต้อง 100% ซึ่งมีความสามารถเข้าใจถึงคำพูดได้ดีไม่ว่าจะเป็นคำพูดของใครก็ตาม เป็นอิสระจากขนาดของกลุ่มคำศัพท์ , ความดัง, ลักษณะของผู้พูด และการออกเสียง หรือเงื่อนไขของช่องทางต่างๆที่เป็นไปได้



ภาพที่ 2. ภาพรวมการทำงานของ Speech Recognition

1.4.3 การตรวจจับการเคลื่อนไหวโดยใช้กล้อง Kinect และ Kinect SDK Beta

ตำแหน่งของข้อต่อที่ได้รับจาก API จะสัมพันธ์กับตำแหน่งของร่างกายมนุษย์ 20 ตำแหน่งในท่ายืน และ 10 ตำแหน่งในท่านั่ง ตามที่แสดงในภาพที่ 2.10 เมื่อผู้ใช้ยืนหันหน้าเข้าหากล้อง ตำแหน่งของข้อต่อข้างซ้ายและขวาจะสลับข้างกัน แตกต่างจากร่างกายจริง แต่ละตำแหน่งของข้อต่อจะมีค่า X, Y และ Z ในหน่วยมิลลิเมตร และมีฟิลด์ Confidence แสดงสถานะ การตรวจจับตำแหน่งข้อต่อนั้นๆ โดยจะมีค่าระหว่าง 0 – 1 ซึ่ง 1 หมายถึง สามารถตรวจจับได้ และหากมีค่าน้อยกว่า 1 (เป็นตัวเลขทศนิยม) หมายถึง ไม่สามารถตรวจจับได้ ข้อมูลตำแหน่งของข้อต่อในส่วนนั้นจะเป็นค่าประมาณ ในบางครั้งอาจเป็นค่าที่ตรวจจับได้ล่าสุดก่อนที่จะตรวจจับไม่ได้ในเวลาต่อมา หรืออาจเป็นค่าที่ไม่มีความหมายก็ได้ ในการตรวจจับการเคลื่อนไหว หากผู้ใช้อยู่ภายในมุมกล้อง ห่างจากกล้องประมาณ 1.2 – 3.5 เมตร (ระยะห่างอุดมคติตามที่ Microsoft แนะนาคือ ประมาณ 2.5 เมตร) ค่าตำแหน่งของข้อต่อที่ ถูกตรวจจับได้ จะมีค่าใกล้เคียงกับการวัดจริง เช่น ความยาวของข้อศอกไปจนถึงกลางฝ่ามือ เป็นต้น และฟิลด์ Confidence ของข้อต่อจะมีค่าเป็น 1 แต่หากมีบางข้อต่อที่กล้องไม่สามารถตรวจจับได้ หรือถูกบดบังไป เช่น ยืนมือไพล่หลัง เป็นต้น ฟิลด์ Confidence ของข้อต่อนั้นจะมีค่าน้อยกว่า 1

1.4.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยเรื่อง **“Evaluating a Dancer's Performance using Kinect-based Skeleton Tracking”** [3] (D. Alexiadis, P. Kelly, P. Daras, N. E. O’Connor, T. Boubekeur, and M. B. Moussa, 2011 : 659 - 662) ได้ทำการศึกษาเกี่ยวกับการวัดประสิทธิภาพของนักเต้นโดยใช้กล้องคีเนคในการตรวจจับการเคลื่อนไหว และใช้ OpenNIเป็น API ในการให้ข้อมูลการตรวจจับ งานวิจัยนี้ใช้การเปรียบเทียบตำแหน่งของข้อต่อของนักเต้นสองคนที่ได้จากการตรวจจับตามเวลาจริงมาเป็นส่วนหนึ่งในการให้คะแนนการเต้น ผลการศึกษาของงานวิจัยนี้แสดงให้เห็นว่าสามารถนำข้อมูลตำแหน่งของข้อต่อและมือ ไปใช้งานในด้านการตรวจจับการเคลื่อนไหวได้

งานวิจัยเรื่อง **“Real Time Extraction of Body Soft Biometric from 3D Videos”** [4] (C. Velardo and J. Dugelay, 2011 : 781-782) ได้ทำการศึกษาการใช้ คีเนคในการใช้ท่าทางและการรู้จำเสียงในการควบคุมส่วนติดต่อกับผู้ใช้บนโปรแกรมกูเกิ้ลเอิร์ธ โดยผลการศึกษาสามารถควบคุมได้ทั้งเสียงและท่าทางผลการศึกษาของงานวิจัยนี้แสดงให้เห็นว่าสามารถนำวิธีการรู้จำเสียงและท่าทางมาใช้ในการใช้เสียงตรวจจับคำพูดได้

งานวิจัยเรื่อง “การประกอบรวมโครงร่างมนุษย์จากการตรวจจับการเคลื่อนไหวโดยใช้กล้อง Kinect หลายตัว” [5] (นราวุฒิ พัฒโนทัย, พรชัย มงคลนาม และ บัณฑิต วรรธนาภา, 2012 : 370-374) ได้ทำการศึกษาการใช้ คีเนคสามตัวต่อเข้าด้วยกันในระบบ Client – Server โดยวางมุม

กล้องทำมุม 120 องศาต่อกัน การศึกษาของงานวิจัยนี้แสดงให้เห็นว่าสามารถเพื่อช่วยเพิ่มโอกาสในการมองเห็นจุด Skeleton ของคนในการจับการเคลื่อนไหวได้ดีขึ้น

ในปี 2012 บริษัท Samsung Electronic[6] ได้นำเทคโนโลยีการรู้จำเสียงและการตรวจสอบการเคลื่อนไหวใส่กับอินเทอร์เน็ตทีวี หลักการก็คล้ายๆกับ คีเนคที่กล้องจะจับท่าทางและตัวทีวีก็คอยจะตอบสนองคำสั่งตามท่าทางของเรา ตัวอย่างเช่น "การแบมือ" โชว์หน้ากล้องก็จะเริ่มต้นคำสั่ง หลังจากนั้นตัว Cursor ลูกศรก็จะปรากฏขึ้นมาบนหน้าจอ เลื่อนมือไปในทิศทางต่างๆ เพื่อเลื่อนตัวลูกศรที่เป็น Cursor ไปในทิศทางที่เราต้องการ หากต้องการกดปุ่มใดๆบนหน้าจอ ก็กำมือ ซึ่งเปรียบเสมือนการคลิกเมาส์ การสั่งงานด้วยเสียง เพียงแค่พูดว่า "HI TV" เพื่อเปิดหรือปิดการใช้งาน Voice Control หลังจากนั้นจะมี "แถบชุดเมนูคำสั่ง" ขึ้นมาด้านล่าง

ชุดคำสั่ง Voice Control

1. "TV Power Off" :ปิดทีวี
2. "Source" :เปลี่ยนแหล่งสัญญาณ
3. "Channel Number" :เลือกช่องทีวีตามที่ต้องการ
4. "Channel Up หรือ Down" :เลื่อนช่องทีวีขึ้นหรือลงทีละช่อง
5. "Volume Up/Down" :เพิ่มระดับเสียงดัง/เบา
6. "Mute" :ปิดเสียง

ดังนั้นงานวิจัยนี้จึงได้เสนอและออกแบบและพัฒนาเทคนิคการผสมผสานเทคนิคระหว่างการรู้จำเสียงและเทคนิคการตรวจสอบการเคลื่อนไหวเข้าด้วยกันเพื่อใช้ในการควบคุมอินเทอร์เน็ตทีวีโดยออกแบบการควบคุมไว้ดังนี้ เปลี่ยนช่อง 3,5,7,9,11,thai PBS เพิ่มลดเสียงได้ 5 ระดับและการเปิดปิดอินเทอร์เน็ตทีวี

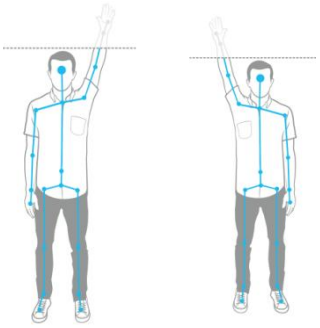
2. วิธีดำเนินการวิจัย

วิธีการผสมผสานระหว่างเทคนิคการรู้จำเสียงและการตรวจสอบการเคลื่อนไหวโดยใช้กล้องคีเนคประกอบไปด้วยขั้นตอนที่มีรายละเอียดดังต่อไปนี้

2.1 การตรวจสอบการรู้จำเสียง

ในการตรวจสอบการรู้จำเสียงอินพุทที่เข้ามาผ่านไมโครโฟนบนตัวของกล้องคีเนคได้นำไลบรารีของ Microsoft Speech Application Programming Interface5.1 (SAPI 5.1) โดยจะเขียนโปรแกรมผ่าน Visual Studio 2010 เพื่อนำอินพุทเสียงที่ได้แปลงเป็นตัวอักษรโดยใช้เทคนิค Speech to Text (STT) จากนั้นนำคำที่ได้ไปเปรียบเทียบกับคำที่ระบุไว้หากไม่ตรงตามที่ระบุไว้ก็จะ Reject ทิ้ง หากตรงตามเงื่อนไขก็จะ Accepted โดยจะมีกระบวนการเปิดและปิดระบบการรู้จำเสียง จะ

กำหนดให้ผู้ใช้ยกมือขวาขึ้นเหนือศีรษะจะเป็นการเปิดระบบการรู้จำเสียงและหากต้องการปิดให้ผู้ใช้ยกมือซ้ายขึ้นเหนือศีรษะจะเป็นการปิดระบบเสียง โดยจะอาศัย Joint Hand Right และ Joint Hand Left ทำการเปรียบเทียบตำแหน่ง Position ด้านแกน Y กับ Joint Head ดังภาพที่3.

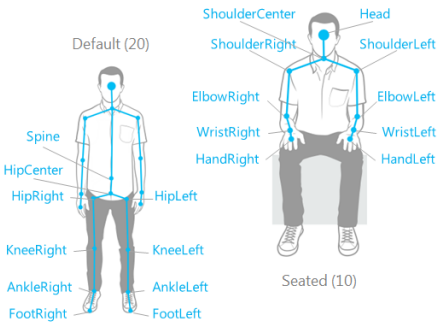


ภาพที่ 3. วิธีการเปิดปิดระบบการสั่งการด้วย

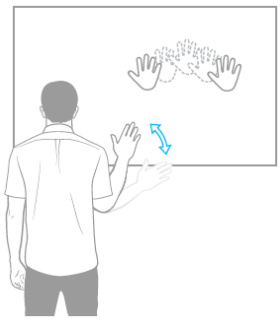
ภาพที่ 4. แสดงรูปภาพระบบการสั่งงานด้วยเสียง

2.2 วิธีการตรวจสอบการเคลื่อนไหว

ในการตรวจสอบการเคลื่อนไหวภาพผ่านหน้ากล้องคีเนคได้นำไลบรารีของ Kinect SDK โดยจะเขียนโปรแกรมผ่านโปรแกรม Visual Studio 2010
โดยทำการตรวจจับรูปร่างของคนแล้วจะนำ Skeleton ไปซ้อนทับบนภาพ โดยท่ายืนจะกำหนดเป็น 20 Joint และท่านั่งจะกำหนดเป็น 10 Joint จากนั้นก็ทำการเขียนโปรแกรมเก็บค่า เฉพาะ Joint Hand Right และ Joint Hand Left เพื่อนำค่า Position ของตำแหน่งมือไปเปรียบเทียบกับตำแหน่ง Cursor ลูกศรบนจอภาพ เพื่อให้ได้ตำแหน่งและพิกัด สำหรับการควบคุมอินเทอร์เน็ตทีวี



ภาพที่ 5. แสดงชื่อ Skeleton Joints



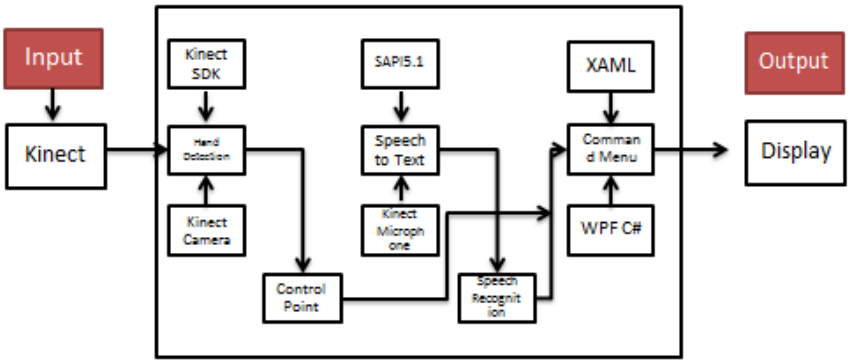
ภาพที่ 6. Cursor ที่ใช้เปรียบเทียบกับ Position

2.3 การผสมผสานวิธีการรู้จำเสียงและการตรวจจับการเคลื่อนไหว

เมื่อได้วิธีการรู้จำเสียงและการตรวจจับการเคลื่อนไหวแล้ว จึงได้ออกแบบวิธีการผสมผสานเทคนิคการรู้จำเสียงและการตรวจสอบการเคลื่อนไหวโดยกำหนดให้ เมื่อ Position ของ Cursor บนหน้าจออินเทอร์เน็ตทีวี ไป Detect กับปุ่มใดๆบนอินเตอร์เฟสบนหน้าพร้อมทั้งออกเสียง

“Yes” ก็จะเป็นการสั่งงานกับเมนูนั้นๆ ดังรูปที่ 1 แต่ถ้าหาก Cursor ไป Detect กับปุ่มใดๆบนอินเตอร์เฟสบนหน้าจอแล้วมีการออกเสียงอื่นโปรแกรมก็จะไม่มี

การทำงาน หรือถ้าออกเสียงอย่างเดียวแล้วไม่ได้มีการนำ Cursor ไป Detect กับปุ่มใดๆ โปรแกรมก็จะไม่มีการทำงานอีกเช่นกัน



ภาพที่ 7. แสดงรูปแบบการทำงานของวิธีการ
ผสมผสานเทคนิคการรู้จำเสียงและการตรวจสอบการ

3. ผลการวิจัย

งานวิจัยนี้ทำการทดสอบประสิทธิภาพโดยการวัดความเร็วและความถูกต้องของวิธีการที่นำเสนอ โดยการพัฒนาโปรแกรมทั้งสามระบบ คือ 1.ระบบการรู้จำเสียงสำหรับการควบคุมอินเทอร์เน็ตทีวี 2.ระบบการตรวจสอบการเคลื่อนไหวสำหรับการควบคุมอินเทอร์เน็ตทีวี 3.ระบบผสมผสานระหว่างการรู้จำเสียงและการตรวจจับการเคลื่อนไหว ขึ้นมาด้วยภาษา C# ใช้ Kinect SDK เวอร์ชัน Beta 1 และสร้างสถานการณ์สภาพแวดล้อมในห้องขึ้นมา โดยให้สภาพแวดล้อมแรกเป็นสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวน โดยที่ค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB สภาพแวดล้อมที่สองคือสภาพแวดล้อมที่มีเสียงรบกวนโดยกำหนดให้เสียงรบกวนอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB ในการตรวจจับตั้งกล้องอยู่บนระนาบที่ขนานกับระนาบพื้นที่ทำการตรวจจับสูงจากระนาบพื้น 0.6 เมตร ห่างออกไป 1.5 เมตร โดยกำหนดโจทย์ที่ใช้ทดสอบดังต่อไปนี้ ปิด TV, เลือกช่อง (เปิด TV), เปลี่ยนช่อง, เพิ่มเสียง, ลดเสียง

ตารางที่ 1. แสดงถึงงานที่ผู้ทดสอบต้องทำในแต่ละชุดการทดสอบ

โจทย์สำหรับใช้ทดสอบ		
ชุดทดสอบ A	ชุดทดสอบ B	ชุดทดสอบ C
เลือกช่อง 3	เลือกช่อง 5	เลือกช่อง 9
เปลี่ยนช่อง 7	เพิ่มเสียงขึ้น 1 ระดับ	เพิ่มเสียงขึ้น 1 ระดับ
ลดเสียงลง 1 ระดับ	เปลี่ยนช่อง 9	ปิด TV
เพิ่มเสียงขึ้น 1 ระดับ	ลดเสียงลง 1 ระดับ	เปลี่ยนช่อง 11
ปิด TV	ปิด TV	ลดเสียงลง 1 ระดับ

โดยจะใช้ผู้ทดสอบจำนวน 60 คนนักศึกษาชั้นปี 1 – 3 ช่วงอายุประมาณ 19 – 21 ปี โดยจะแบ่งการทดสอบเป็น 2 กลุ่ม กลุ่มล่ะ 30 คนโดยให้กลุ่มแรกทำการทดสอบในสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวนค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB และกลุ่มที่สองอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB โดยให้แต่กลุ่มทำการทดสอบดังนี้

1.ให้ผู้ทดสอบ สุ่มเลือก เทคนิคสำหรับควบคุมอินเทอร์เน็ตทีวี จาก 3 วิธี มา 1 วิธี พร้อมทั้งนำหัวข้อเทคนิคที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

2.จากนั้นให้ผู้ทดสอบ สุ่มเลือก ชุดทดสอบ จาก 3 ชุดทดสอบ มา 1 ชุดทดสอบ พร้อมทั้งนำหัวข้อชุดทดสอบ ที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

3.ให้ผู้ทดสอบ เข้าอยู่ในบริเวณที่จะทดสอบ หากผู้ทดสอบเตรียมตัวพร้อมผู้ประเมินจะเริ่มทำการจับเวลาในการทดสอบ

4.เมื่อผู้ทดสอบ ทำตามเงื่อนไขจากชุดทดสอบจนเสร็จ ผู้ประเมินจะทำการหยุดเวลาพร้อมจดบันทึกเวลาที่ผู้ทดสอบได้ใช้ไป

5.ให้ผู้ทดสอบทำการ สุ่มเลือก เทคนิคสำหรับควบคุมอินเทอร์เน็ตทีวี จาก 2 วิธีที่เหลือมา 1 วิธี พร้อมทั้งนำหัวข้อเทคนิคที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

6.จากนั้นให้ผู้ทดสอง สุ่มเลือก ชุดทดสอบ จาก 2 ชุดทดสอบ มา 1 ชุดทดสอบ พร้อมทั้งนำหัวข้อชุดทดสอบ ที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

7.ให้ผู้ทดสอบ เข้าอยู่ในบริเวณที่จะทดสอบ หากผู้ทดสอบเตรียมตัวพร้อมผู้ประเมินจะเริ่มทำการจับเวลาในการทดสอบ

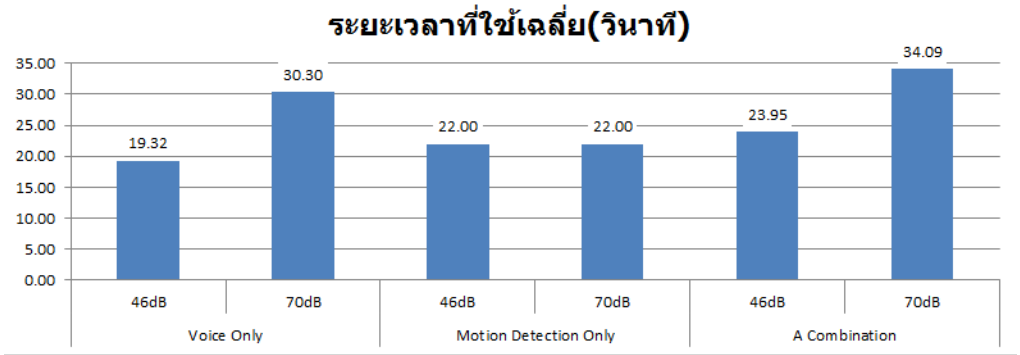
8.เมื่อผู้ทดสอบ ทำตามเงื่อนไขจากชุดทดสอบจนเสร็จ ผู้ประเมินจะทำการหยุดเวลาพร้อมจดบันทึกเวลาที่ผู้ทดสอบได้ใช้ไป

9.จากนั้นให้ผู้ทดสอบทำการทดสอบ เทคนิคสำหรับควบคุมอินเทอร์เน็ตทีวีและชุดทดสอบ ที่เหลืออยู่

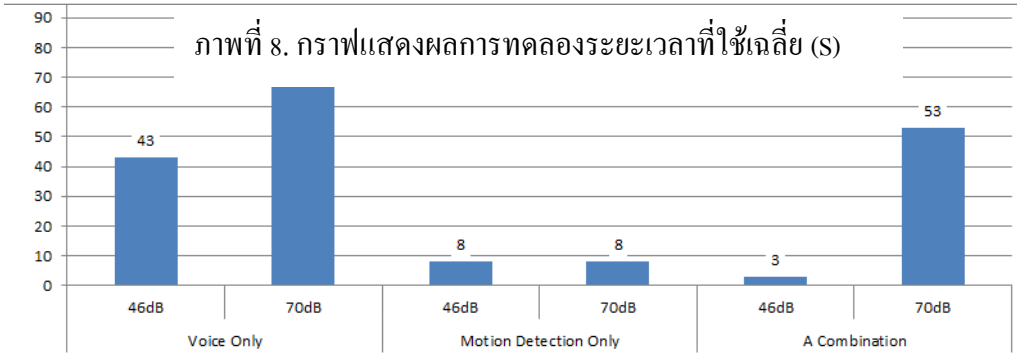
10.ให้ผู้ทดสอบ เข้าอยู่ในบริเวณที่จะทดสอบ หากผู้ทดสอบเตรียมตัวพร้อมผู้ประเมินจะเริ่มทำการจับเวลาในการทดสอบ

11.เมื่อผู้ทดสอบ ทำตามเงื่อนไขจากชุดทดสอบจนเสร็จ ผู้ประเมินจะทำการหยุดเวลาพร้อมจดบันทึกเวลาที่ผู้ทดสอบได้ใช้ไป

12.นำข้อมูลที่ได้จากทั้ง 3 กลุ่ม มาวิเคราะห์เปรียบเทียบความแตกต่างในเรื่องของความเร็วและความถูกต้องของการทำงาน



ภาพที่ 8. กราฟแสดงผลการทดลองระยะเวลาที่ใช้เฉลี่ย (S)



ภาพที่ 9. กราฟแสดงผลการทดลองจำนวนครั้งที่ผิดพลาดในการ

จากผลงานทดลองส่วนวิเคราะห์การใช้วิธีรู้จำเสียงในการสั่งงานอินเทอร์เน็ตทีวีพบว่าสามารถสั่งงานได้รวดเร็วแต่จะมีความผิดพลาดค่อนข้างมากทั้งสองสถานการณ์เนื่องจากคำศัพท์ที่ใช้ในการควบคุมมีปริมาณของคำสั่งมากทำให้ความเป็นไปได้ของเสียงที่จะออกถูกต้องมีน้อยลง อีกทั้งหากเวลาควบคุมมีเสียงจากทีวีเข้าสู่ตัวรับเสียงก็ยิ่งทำให้เกิดอัตราการผิดพลาดของคำสั่งมากขึ้นไปอีกด้วย อาจแก้ไขด้วยวิธีการย้ายอุปกรณ์รับเสียงให้อยู่ใกล้ตัวผู้สั่งการมากที่สุด

เพื่อที่จะรับเสียงได้อย่างถูกต้องชัดเจนและเสียงรบกวนในห้องน้อยที่สุดเพื่อเพิ่มความถูกต้องของการทำงาน

ส่วนวิเคราะห์การตรวจสอบการเคลื่อนไหว นั้นสามารถทำได้ดีในทั้งสถานการณ์ทั้งสอง เนื่องจากเสียงรบกวนไม่มีผลกระทบต่อเทคนิคนี้ และทำงานได้อย่างถูกต้อง แต่จะทำงานได้ค่อนข้างช้า เนื่องจากผู้สั่งการต้องอยู่ในระยะที่กล้องสามารถตรวจจับได้และแสงสว่างภายในห้องต้องเพียงพอต่อการมองเห็นของกล้อง ก็จะทำให้การทำงานถูกต้องมากยิ่งขึ้น

ส่วนวิเคราะห์การผสมผสานระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว จะพบว่าในสถานการณ์ที่หนึ่งคือเสียงในห้องอยู่ในระดับ 46 dB การใช้วิธีนี้จะได้ผลการทำงานที่ถูกต้องมากที่สุดเนื่องจากตัวโปรแกรมต้องยืนยันการทำงานถึงสองกระบวนการคือตรวจสอบจากเคลื่อนไหวของมือเมื่อการเคลื่อนที่ของมือถูกต้องจึงมาตรวจสอบด้วยเสียงอีกครั้งหนึ่งข้อดีของวิธีนี้เราจะมีคำที่ใช้ยืนยันแค่คำเดียวคือ Yes เพื่อลดความผิดพลาดในเรื่องของคำหรือประโยคที่คล้ายๆกันในการสั่งการอินเทอร์เน็ตทีวี ทำให้การทำงานออกมาค่อนข้างถูกต้องสมบรูณ์ มากที่สุดในกรณีที่หนึ่ง ส่วนในกรณีที่สองคือเสียงภายในห้องเฉลี่ยนอยู่ที่ 70 dB การทำงานเริ่มจะมีความผิดพลาดเนื่องจากมีเสียงรบกวนจากภายนอกทำให้ตัวโปรแกรมไม่สามารถตรวจจับคำได้ แต่อย่างไรก็ดีวิธีนี้โปรแกรมก็ยังสามารถตรวจจับคำได้ดีกว่าการใช้วิธีการรู้จำเสียงอย่างเดียวเพราะโปรแกรมจะสามารถจับคำหรือประโยคสั้นๆได้ดีกว่าประโยคยาวๆที่มีหลากหลายของคำสั่ง

4. อภิปรายผล

งานวิจัยชิ้นนี้ได้นำเสนอเทคนิคสำหรับการควบคุมอินเทอร์เน็ตทีวีโดยคีเนคด้วยการผสมผสานระหว่างการรู้จำเสียงและการตรวจสอบการเคลื่อนไหวจากผลของวิธีการที่นำเสนอแสดงให้เห็นว่า ผลลัพธ์ที่ได้ในเรื่องของประสิทธิภาพของความถูกต้องในการควบคุมอินเทอร์เน็ตทีวีมีความถูกต้องมากที่สุดในสถานการณ์ที่มีเสียงรบกวนน้อย คืออยู่ที่ 46 dB เนื่องจากเป็นการจรวจสอบถึงสองขั้นตอนวิธี ทั้งการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว แต่ก็มีข้อเสียในเรื่องของความเร็ว อาจเป็นเพราะผู้ทดสอบเพิ่งเคยได้ใช้เทคโนโลยีที่ไม่เคยได้ถูกการฝึกมาก่อน แต่เพื่อประโยชน์อย่างแท้จริงในอนาคตควรที่จะมีการทดลองในรูปแบบต่างๆมากขึ้นเช่น กลุ่มทดลองกลุ่มที่หนึ่งที่ได้รับการฝึกฝนการใช้กับอีกกลุ่มที่ไม่ได้รับการฝึกฝนการใช้งานเพื่อวัดผลทางด้านประสิทธิภาพและประสิทธิผลของแต่ล่ะวิธี

บรรณานุกรม

[1] Microsoft. **Kinect for Xbox 360 - Xbox.com**. [Online]. Available: www.xbox.com/en-GB/kinect, 2012.

[2] Microsoft. **Speech recognition – Sapi5.1**. [Online]. Available:www.microsoft.com/en-us/download/details.asp ?id=10121

[3] D. Alexiadis, P. Kelly, P. Daras, N. E. O'Connor, T. Boubekeur, and M. B. Moussa, "Evaluating a Dancer's Performance using Kinect-based Skeleton Tracking"**Proceedings of the 19th ACM international conference on Multimedia**. pp. 659-662:Nov. 2011.

[4] Velardo, C. and Dugelay, J., 2011, "Real Time Extraction of Body Soft Biometric from 3D Videos", Proceedings of the 19th ACM international conference on Multimedia, 28 November-1 December, Arizona, USA, pp. 781-782.

[5] นราวุฒิ พัฒโนทัย. พรชัย มงคลนาม. บัณฑิต วรรธนาภา."การประกอบรวมโครงร่างมนุษย์จากการตรวจจับการเคลื่อนไหวโดยใช้กล้อง Kinect หลายตัว" **National Conference on Information Technology 2012**. 370-374 : April. 2012.

[6] Roman and Boomerang.**รีวิว Samsung Smart TV 55ES8000**. [Online]. Available:http://www.lcdtvthailand.com/review/detail.asp?desc=1&param_id=1277

ภาคผนวก จ

หนังสือรับรองผลงานวิจัยในการประชุมวิชาการ การบริหารและการจัดการครั้งที่ 8 “การบริหารจัดการแห่งโลกพลวัต” โดยมหาวิทยาลัยธุรกิจบัณฑิตย์ร่วมกับเครือข่ายวิจัยประชาชื่น

ภาพที่ 9 ประกาศนียบัตรรับรองในการนำเสนอผลงานวิชาการ

รายชื่อผู้ทรงคุณวุฒิพิจารณาผลงานวิชาการ

ศาสตราจารย์ ดร.ไพฑูรย์ สินลารัตน์
รองศาสตราจารย์ ดร.เชาว์ โรจน์แสง
รองศาสตราจารย์ ดร.นรา สมประสงค์
รองศาสตราจารย์ ดร.ประดินันท์ อุปรมัย
รองศาสตราจารย์ ดร.สรชัย พิศาลบุตร
รองศาสตราจารย์ ดร.สุชาติ ณ หนองคาย
รองศาสตราจารย์ ดร.อุษา บิ้กกินส์
รองศาสตราจารย์ พินิจ ทิพย์มณี
รองศาสตราจารย์ ยุทธนา ธรรมเจริญ
รองศาสตราจารย์ สมประสงค์ วิทยเกียรติ
รองศาสตราจารย์ องอาจ ปทะวานิช
ผู้ช่วยศาสตราจารย์ ดร.เพียรศิริ เอกนิยม
ผู้ช่วยศาสตราจารย์ ดร.ฐิติมา สิทธิพงษ์พานิพานิช
ผู้ช่วยศาสตราจารย์ ดร.นภาภรณ์ เชี่ยวชาญ
ผู้ช่วยศาสตราจารย์ ดร.ลีลา เตี้ยงสูงเนิน
ผู้ช่วยศาสตราจารย์ ดร.วรสิทธิ์ ชูชัยวัฒนา
ผู้ช่วยศาสตราจารย์ ดร.วราพร จิระพันธุ์ทอง
ผู้ช่วยศาสตราจารย์ ดร.วันวร จะนู
ผู้ช่วยศาสตราจารย์ ดร.อดิลล่า พงศ์ยี่หล้า
ผู้ช่วยศาสตราจารย์ ปิยวรรณ แสงสว่าง
ผู้ช่วยศาสตราจารย์ ผ่องใส เพ็ชรรักษ์
ผู้ช่วยศาสตราจารย์ สุเทพ พันประสิทธิ์
ผู้ช่วยศาสตราจารย์ เฉลิมพร อภิชนาพงศ์
ผู้ช่วยศาสตราจารย์ เชียง เภาชิต
ผู้ช่วยศาสตราจารย์ เลิศพร ภาระสกุล
ผู้ช่วยศาสตราจารย์ จุลศักดิ์ ชาญณรงค์
ดร.เกียรติอนันต์ ล้วนแก้ว
ดร.เยาวลักษณ์ ราชแพทยาคม
ดร.เอกพงศ์ กิตติสาร
ดร.ธัญธัช วิภัติภูมิประเทศ
ดร.ปริญ เฟื่องวุฒิ
ดร.รชฎ ขำบุญ
ดร.สาวิตรี สมุทรจักร
ดร.อรวรรณ อิ่มสมบัติ
อาจารย์วิญญู นิรนาทล้ำพงศ์