Hand gesture recognition using Skin Likelihood Image

Abstract:

Гарны дохио зангааг таних хэд хэдэн аргууд () зэрэг судалгааны ажилуудад дэвшигдсэн. Эдгээр ажилууд ерөнхийдөө 4-н хэсгээс бүрдэнэ (Hand detection, Hand tracking, Hand Segmentation, Gesture recognition). Энэхүү судалгааны ажилаар дээрхи хэсгүүдэд skin likelihood map-ыг ашиглан хурд болон найдвартай байдлыг хэрхэн ихэхгэж болох талаар судалсан

Intro:  
1. Hand Detection

1. Viola and Jones’s Cascaded Classifier (VJ): Одоогийн байдлаар обьект таних хамгийн шилдэг арга. Харьцангуй хурдан бөгөөд өнгөнөөс үл хамааран обьектыг танидаг. Гар зэрэг хэлбэр дүрс өөрчилөгддөг динамик дүрсийн хувьд False Positive-нь хэт ихэсдэг.
2. Skin-Map based Blob detection: Арьсны pixel бүлэглэдэг бөгөөд хялбар боловч нэмэлтээр depth-map ашиглан гарыг танидаг.
3. Face-Detection and Skin-Map based Hand detection: Өмнөх аргад Depth Map ашиглан гарын хэсгийг ялгасан бол, энэ удаа VJ ашиглан олсон царайны хэсгийг зургаас хасаж гарыг танина.

2. Hand Tracking

1. MeanShift болон CamShift: Хурдан бөгөөд хялбар боловч, Арын дэвсгэрийн өнгөнөөс хэт хамааралтай
2. Flocks of Features: KLT, Edge зэрэг зурагны featue-үүдийн бүлгийг ашиглан дагадаг. Үр дүн нь CamShift-ээс ойролцоогоор 2-дахин хурдан боловч харьцангуй удаан.

3. Hand segmentation

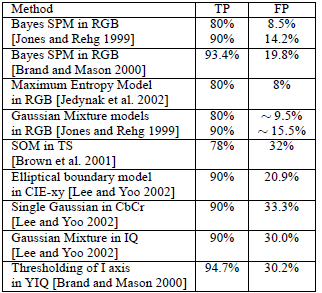
1. Ихэвчлэн Detection болон Tracking хийж олсон Region Of Interest-ийн тодорхой нэгж хэсгээс (Жишээ нь алганы хэсэг) авсан Histogram-ыг ашиглан Background-оос ялган авдаг

4. Gesture Recognition-руу арай ороогүй байгаа

Proposed Method:

1. Skin map

Арьсны өнгийг таних хамгийн тохиромжтой арга нь [] Хүснэгт (1)-ээс харвал 32x32x32 хэмжээс дээрхи RGB histogram дээр тулгуурласан Bayes загвар нь хамгийн өндөр үр дүн үзүүлсэн. Харин Skin Map-ийг таних Decision Tree нь маш бага False Positive-тэй байсан ч гэрэлтүүгт хэт мэдрэг байсан Зураг ().



Хүснэгт(1)

Иймд энэхүү 2 аргыг дараах томьёг ашиглан хольсон үр дүн Зураг ().

falseFactor-нь 1-ээс бага бутархай тоо, Iij нь үндсэн зурагын pixel



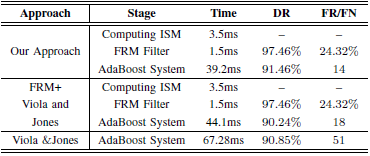
Decision Tree Bayes



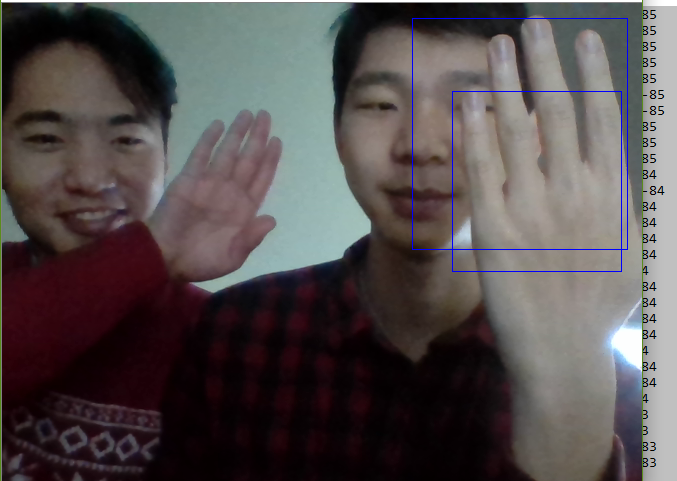
Bayes + Decision Tree (falseFactor=0.5)

2. VJ

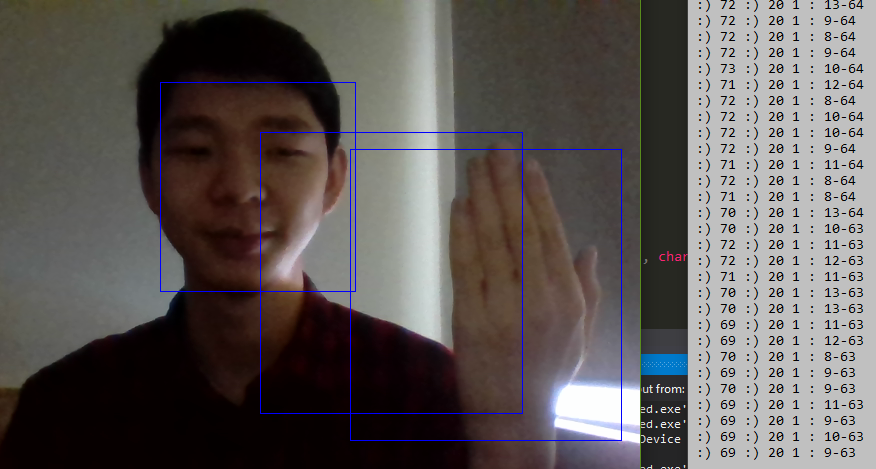
[] ажилд VJ-ийн алгоритмд Skin Map ашиглан False Positive Rate-ийг багасгаад зогсохгүй үндсэн алгоритмийн хурдыг нэмсэн Хүснэгт(2). Уг аргын нэг дутагдалтай тал нь Skin Integral Image болон GrayScale Integral Image 2-ийг зэрэг бодох байсан тул шууд болон дан Skin Integral Image ашиглан гар таних боломжтой эсэхийг судалсан.



Хүснэгт(2)



Normal VJ With Skin filter



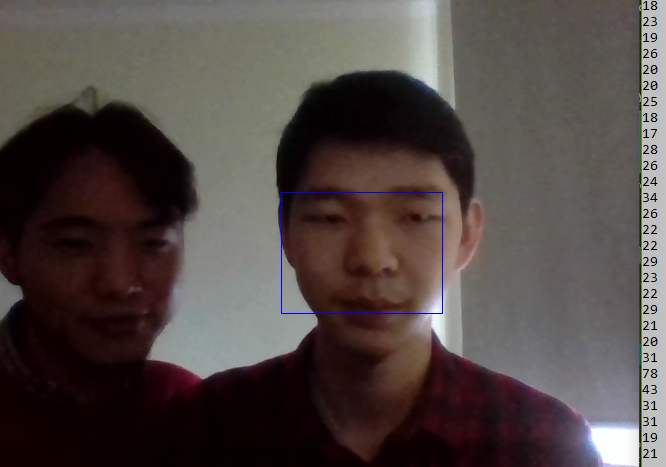
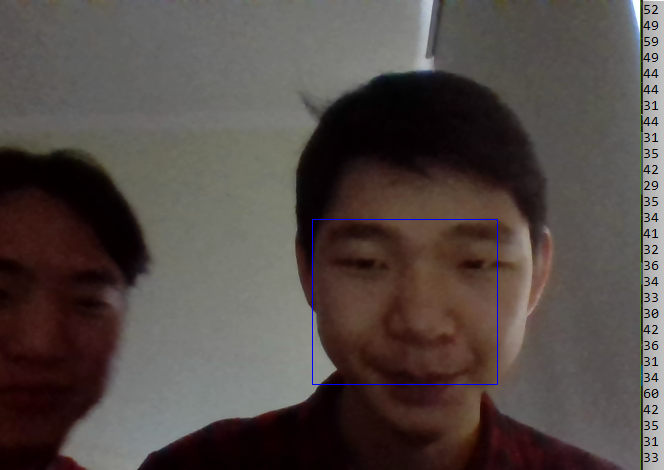
Bayes VJ

Дээрхи зурагнуудаас үзэхэд Skin Integral Image дангаар ашиглах нь False Positive- нь хэт ихэссэн. Алгоритмийн хурд 84ms ээс 44ms болсон ба нийт зургын 60%-ийг алгассан.

3. MeanShift

MeanShift аргыг дурын Probability Density Function-д ашиглаж болох бөгөөд HSV Scale-ийн Hue талбар дээрхи дагах обьектийн магадлалыг ихэвчлэн ашигладаг. Гэхдээ ингэснээр дэвсгэрийн өнгө болон орчны гэрэлд хэт мэдрэг болох бөгөөд зурагны pixel болгонд hue утгыг бодох хэрэгтэй болно. Hue утгыг бодох аргыг зураг№ -д харуулав.

Харж байгаачлан Hue утгыг бодох нь харьцангуй computationally expensive тул арьсны өнгө биш хэсэг буюу Skin Likelihood Image-ийн тодорхой threshold-оос бага утгуудыг бодолгүй орхивол ихээхэн цаг хэмнэнэ. Мөн илүү дутуу өнгөний мэдээллийг хассан тул илүү нарийвчлалтай болж гэрэлд мэдрэг байдал нь багасна. Дараахь зургуудад үр дүн болон магадлалын тархалтыг харуулав.



Original Meanshift With Skin Map

Tracking хийх дундаж хугацаа нь 44ms ээс 25ms болсон.

Дутуу зүйлс:

**Threshold тохируулах** : Дээрхи хоёр аргад ямар хэмжээний утгатай pixel эсвэл цонхыг алгасах вэ гэдгийг тодорхойлох Threshold-ыг тохируулах хэрэгтэй. Гэхдээ уг утга нь гэрэлтүүлгээс их хамаарч байгаа тул түүнийг тодорхойлох функцыг загварчилах.

**Hand Segmentation** : Mean Shift-д бодсон Histogram-ыг ашиглан тодорхой threshold ашиглан гарын хэсгийг ялгах бодолтой байгаа.

**Gesture Recognition :**

**Main Algorithm:**