



UNIVERSITÀ DEGLI STUDI ROMA TRE

Dipartimento di Ingegneria

---

Corso Di Laurea Triennale In

INGEGNERIA INFORMATICA

VALUTAZIONE DELLA QUALITY OF  
EXPERIENCE TRAMITE SENSORE KINECT

Tesi di Laurea di

VALERIO CETORELLI

**Relatore**

Federica Battisti

Anno Accademico 2014/2015



*Alla mia famiglia,  
soprattutto a quel mio fratello sempre sotto pressione. Allons-y!*



# Indice

<b>INTRODUZIONE.....</b>	<b>1</b>
<b>1    RICONOSCIMENTO DELLE EMOZIONI.....</b>	<b>3</b>
1.1.    RICONOSCIMENTO FACCIALE.....	3
1.2.    RICONOSCIMENTO DELLE EMOZIONI.....	4
1.2.1. <i>Paul Ekman ed i suoi studi sulle emozioni</i> .....	5
1.3.    MIMICA FACCIALE.....	5
1.3.1. <i>Facial Action Coding System e Action Units</i> .....	6
1.4.    EMOZIONI PRIMARIE.....	7
1.4.1. <i>Sorpresa</i> .....	8
1.4.2. <i>Gioia</i> .....	8
1.4.3. <i>Paura, disgusto e rabbia</i> .....	9
1.4.4. <i>Tristezza</i> .....	10
1.5.    SCELTA DELLE EMOZIONI RILEVATE.....	11
<b>2    KINECT.....</b>	<b>13</b>
2.1.    COMPONENTI HARDWARE.....	13
2.2.    TECNICHE PER L'ACQUISIZIONE DI IMMAGINI 3D.....	14
2.3.    KINECT FOR WINDOWS.....	17
2.4.    SDK UTILIZZATI.....	18
2.4.1. <i>Kinect For Windows SDK 1.8</i> .....	18
2.4.2. <i>Face Tracking SDK</i> .....	19
2.4.2.1. <i>Animation Units</i> .....	20
2.5.    PERCHÉ SI È SCELTA LA KINECT.....	22
<b>3    IMPLEMENTAZIONE.....</b>	<b>23</b>
3.1.    WINDOWS PRESENTATION FOUNDATION.....	23
3.2.    PANORAMICA DELL'APPLICAZIONE.....	24
3.3.    AGGIUNTA DEL METODO GETSHAPEPOINTS.....	25
3.4.    INTERFACCIA GRAFICA.....	27
3.5.    EMOTIONSANALIZER.....	28
3.5.1. <i>Rilevazione attraverso Animation Units</i> .....	30
3.5.1.1. <i>Test eseguiti sulle Animation Units</i> .....	30
3.5.1.2. <i>Margini di riferimento</i> .....	31
3.5.1.3. <i>Animation Units utilizzate</i> .....	32
3.5.2. <i>Scelta per la rilevazione del disgusto</i> .....	32
3.5.3. <i>Setup dell'applicazione</i> .....	33
3.5.4. <i>Metodo AnalyzeEmotions</i> .....	34

<b>4</b>	<b>TEST .....</b>	<b>36</b>
4.1.	SETUP SPERIMENTALE .....	36
4.2.	SCELTA DEI SOGGETTI .....	37
4.3.	FASE DI TRAINING.....	37
4.4.	PARAMETRI DI VALUTAZIONE.....	38
4.5.	PRIMA SESSIONE DI TEST .....	39
4.5.1.	<i>Fase di rilevazione</i> .....	39
4.5.2.	<i>Elaborazione dei risultati</i> .....	39
4.5.3.	<i>Analisi dei risultati</i> .....	41
4.5.4.	<i>Analisi prestazioni</i> .....	41
4.6.	SECONDA SESSIONE DI TEST.....	42
4.6.1.	<i>Scelta degli stimoli</i> .....	43
4.6.2.	<i>Fase di rilevazione</i> .....	43
4.6.3.	<i>Elaborazione dei risultati</i> .....	44
4.6.4.	<i>Analisi dei risultati</i> .....	44
	<b>CONCLUSIONE E SVILUPPI FUTURI .....</b>	<b>46</b>
	<b>BIBLIOGRAFIA .....</b>	<b>48</b>

## Indice delle Figure

FIGURA 1 MIMICA FACCIALE DELL'ESPRESSIONE NEUTRA [11].....	6
FIGURA 2 ESEMPI DI ACTION UNIT PRESENTI NEL COHN-KANADE DATABASE [14] .....	7
FIGURA 3 MIMICA FACCIALE DELLA SORPRESA [11].....	8
FIGURA 4 MIMICA FACCIALE DELLA GIOIA [16].....	9
FIGURA 5 MIMICA FACCIALE DELLA PAURA [16].....	10
FIGURA 6 MIMICA FACCIALE DEL DISGUSTO [17] .....	10
FIGURA 7 MIMICA FACCIALE DELLA RABBIA [16].....	10
FIGURA 8 MIMICA FACCIALE DELLA TRISTEZZA [18].....	11
FIGURA 9 COMPONENTI DEL SENSORE KINECT .....	14
FIGURA 10 ESEMPIO DI PROIEZIONE DI LINEE.....	15
FIGURA 11 ESEMPIO DI PROIEZIONE DI UN PATTERN PERIODICO .....	15
FIGURA 12 ESEMPIO DI PROIEZIONE DI UN PATTERN PSEUDO-CASUALE [19] .....	16
FIGURA 13 RAGGIO D'AZIONE DEL SENSORE KINECT FOR WINDOWS [25].....	17
FIGURA 14 RAGGIO D'AZIONE DEL SENSORE KINECT FOR WINDOWS CON <i>NEAR MODE</i> [25].....	18
FIGURA 15 MOVIMENTI DESCRITTI DAGLI <i>HEAD POSE ANGLES</i> [28] .....	19
FIGURA 16 POSIZIONE DI ALCUNI DEI 100 <i>TRACKED POINTS</i> [28].....	20
FIGURA 17 ESEMPIO DI PROIEZIONE DEL MODELLO CANDIDE-3 [35] .....	26
FIGURA 18 ESEMPIO DI PROIEZIONE DEI <i>TRACKED POINTS</i> .....	27
FIGURA 19 SCHERMATA PRINCIPALE DELL'APPLICAZIONE.....	28
FIGURA 20 10 ELENCO COMPLETO DEI <i>TRACKED POINTS</i> CON RELATIVO NUMERO [35] .....	33
FIGURA 21 ESEMPIO DI SETUP SPERIMENTALE UTILIZZATO NELLA PRIMA SESSIONE DI TEST .....	37
FIGURA 22 ESEMPIO DI UN FILE DI LOG.....	40
FIGURA 23 GRAFICO CON LE PERCENTUALI DI TP, FN E FP PER LA PRIMA SESSIONE DI TEST .....	41

## Indice delle Tabelle

TABELLA 1 CLASSIFICAZIONE DELLE ANIMATION UNIT [26].....	21
TABELLA 2 ELENCO DELLE CLASSI DELL'APPLICAZIONE.....	24
TABELLA 3 ELENCO DELLE PAGINE XAML DELL'APPLICAZIONE .....	24
TABELLA 4 METODO GETSHAPEPOINTS .....	26
TABELLA 5 LIBRERIE UTILIZZATE NELLA CLASSE <i>EMOTIONSANALIZER</i> .....	28
TABELLA 6 LISTE USATE PER IL SETUP DELL'APPLICAZIONE .....	29
TABELLA 7 VARIABILI RAPPRESENTANTI I VALORI BASE PER LA RILEVAZIONE .....	29
TABELLA 8 METODI PRINCIPALI DELLA CLASSE <i>EMOTIONSANALIZER</i> .....	29
TABELLA 9 RILEVAZIONI ESEGUITE PER L'AU <i>LIP STRETCHER</i> .....	31
TABELLA 10 RILEVAZIONI ESEGUITE PER L'AU <i>OUTER BROW RAISER</i> .....	31
TABELLA 11 MARGINI DI RIFERIMENTO.....	31
TABELLA 12 METODO SETUP .....	34
TABELLA 13 METODO ANALIZEEMOTION .....	35
TABELLA 14 EQUAZIONI PER I PARAMETRI <i>ACCURACY</i> , <i>PPV</i> E <i>NPV</i> .....	38
TABELLA 15 RISULTATI DELLA PRIMA SESSIONE DI TEST.....	41
TABELLA 16 RISULTATI DELLA PRIMA SESSIONE DI TEST SENZA TENER CONTO DI SOGGETTI CON OCCHIALI .....	42
TABELLA 17 RISULTATI DELLA SECONDA SESSIONE DI TEST.....	45





## Introduzione

Con *Quality of Experience* si intende la misura del grado di soddisfazione di un utente, in risposta ad un certo servizio. Solitamente questo grado di soddisfazione è basato sulle risposte che l'utente stesso fornisce ad un questionario specifico.

Le informazioni ottenute da questo tipo di sondaggio non sono però sempre efficaci al 100%: l'utente non sempre è incline alla compilazione di questionari e non sempre è sincero nel fornire le risposte.

Nella trattazione si vogliono quindi fornire le basi per sviluppare un sistema *user friendly*, che non importuni l'utente con una lunga serie di domande e che sia in grado di rilevare con certezza il grado di soddisfazione espresso, attraverso la rilevazione dell'emozione che l'utente stesso ha in risposta ad un dato servizio o contenuto.

La psicologia ha fornito nel corso dell'ultimo secolo diverse tecniche per capire l'emozione di una persona ed è compito di questa trattazione fornire un sistema che riesca a rilevare l'emozione dell'utente in maniera automatica ed in tempo reale, attraverso l'utilizzo del sensore Kinect, rilasciato da Microsoft Corporation.

Il presente lavoro di tesi si compone di quattro capitoli.

Nel primo capitolo verranno introdotti gli studi effettuati riguardanti le emozioni e le relative espressioni sul volto.

Una particolare attenzione verrà riservata alla classificazione delle emozioni e alla loro universalità per tutto il genere umano, facendo riferimento agli studi condotti dallo psicologo Paul Ekman in questo campo.

Nel secondo capitolo verrà preso in esame il sensore Kinect, presentandone le diverse versioni rilasciate e illustrandone il funzionamento delle componenti hardware, con particolare attenzione alle tecniche per il rilevamento di un'immagine 3D.

Verrà spiegato il motivo della scelta di una particolare versione del sensore, anche il relazione alle librerie software fornite, di cui si offre una panoramica nel capitolo stesso.

Il terzo capitolo è dedicato all'implementazione di un programma per il riconoscimento delle emozioni. Vengono descritti i passi fondamentali ed i metodi più importanti del programma sviluppato.

Particolare attenzione viene riservata alla scelta di utilizzare alcune funzionalità delle librerie piuttosto che altre, anche in relazione a test svolti sul funzionamento e sulla precisione delle librerie stesse.

Nel quarto capitolo vengono descritte due sessioni di test svolte per analizzare le prestazioni del programma sviluppato.

La prima sessione si compone di test atti a misurare la precisione del programma in relazione ad altri software già rilasciati al momento della trattazione.

La seconda sessione invece è composta da un test soggettivo con lo scopo di misurare gli stessi parametri del precedente test in una situazione più reale, ai fini di misurare l'andamento dei risultati in entrambi i casi.

# 1 Riconoscimento delle emozioni

---

*Emotions determine the quality of our lives. They occur in every relationship we care about — in the workplace, in our friendships, in dealings with family members, and in our most intimate relationships.*

---

*P. EKMAN - EMOTIONS REVEALED: RECOGNIZING FACES AND FEELINGS TO IMPROVE COMMUNICATION AND EMOTIONAL LIFE*

In questo capitolo viene esaminato lo stato dell'arte riguardante il riconoscimento facciale e il riconoscimento delle emozioni.

## 1.1. Riconoscimento facciale

Il riconoscimento facciale è una branca della biometria che si occupa di stabilire l'identità di una persona a partire dalle caratteristiche del volto.

Con il costante sviluppo della tecnologia il riconoscimento facciale automatizzato è divenuto un obiettivo importante nell'ambito della ricerca. Sono ormai moltissimi i software in circolazione che implementano tale tecnologia con ottimi risultati.

Ognuno dei maggiori attori nell'ambito informatico, al momento della trattazione, ha sviluppato un software con queste finalità.

Di seguito sono riportati alcuni esempi, senza pretese di completezza:

- Google: ha sviluppato un sistema in grado di riconoscere le facce delle persone catturate durante le riprese per le mappe di Google Street View, così da poterle offuscare per motivi di privacy [1];
- Microsoft: permette l'accesso al proprio sistema operativo grazie ad una tecnologia di riconoscimento facciale, in grado di distinguere due gemelli omozigoti [2];

- Facebook: utilizza il riconoscimento facciale per identificare una persona presente in una foto, confrontandone i parametri del volto con quelli presenti in altre foto già inserite nel sistema [3].

L'informatica non è però l'unico settore a beneficiare dei progressi fatti nel riconoscimento facciale: nel campo della sicurezza sono stati sviluppati, ad esempio, Bancomat in grado di riconoscere l'utente, bloccando così eventuali tentativi di frode [4].

## **1.2. Riconoscimento delle emozioni**

Appurato il successo avuto dal riconoscimento facciale in campo informatico, il naturale passo successivo è stato quello di estrapolare informazioni dal volto di una persona per poterne riconoscere così le emozioni.

Le applicazioni di questa tecnologia sono molteplici: un sistema in grado di riconoscere l'emozione di un utente è soprattutto in grado di valutarne l'esperienza in tempo reale, mentre usufruisce di contenuti o servizi, ovvero di valutarne la *Quality of Experience* [5].

Un ambito di ricerca molto importante all'infuori del settore informatico, che sta riscontrando un discreto interesse, mira a riconoscere la rabbia in un automobilista, così da poter riprodurre messaggi per calmarlo, in modo da prevenire incidenti [6].

Il sistema SPOT (*Screening of Passengers by Observation Techniques*), utilizzato in 161 aeroporti statunitensi, con il compito di individuare atteggiamenti sospetti tra i passeggeri ed evitare così comportamenti pericolosi, è invece un chiaro esempio di utilizzo del riconoscimento delle emozioni in ambito di sicurezza e prevenzione [7].

Degno di nota è anche il software *FaceTracker 6.0*, sviluppato dalla compagnia olandese Noldus Information Technology, considerato il miglior software per il riconoscimento delle espressioni al momento della trattazione [7]. Le prestazioni e l'affidabilità di questo software verranno riprese nello specifico nel capitolo relativo ai test svolti, come metro di paragone per il software sviluppato nell'ambito del progetto.

### **1.2.1. Paul Ekman ed i suoi studi sulle emozioni**

Il lavoro sul riconoscimento delle emozioni e delle espressioni facciali portato avanti dallo psicologo statunitense Paul Ekman è considerato un punto di riferimento per ogni ricerca in questi ambiti. Per questo motivo Ekman è considerato dalla comunità scientifica uno dei più importanti psicologi di tutto il ventesimo secolo e una delle 100 persone più influenti del mondo secondo la rivista “*Times Magazine*” [8].

Inizialmente supportato dall’*Advanced Research Projects Agency* (ARPA) del Dipartimento della Difesa statunitense, con lo scopo di aiutare l’attività di controspionaggio, ha condotto studi approfonditi sul riconoscimento della falsificazione di un’emozione, per poter riconoscere un’eventuale menzogna [9], meritando l’appellativo di “*the Best Human Lie Detector in the World*” [10].

Ai fini della trattazione non verranno presi in considerazione i suoi studi in questa direzione, ma ci si focalizzerà sul suo lavoro preliminare, atto a riconoscere e catalogare le diverse espressioni ed emozioni esistenti.

## **1.3. Mimica facciale**

La mimica facciale è una branca della cinesica (la scienza che studia i movimenti del corpo) che si occupa dello studio del volto umano.

In questo ambito due antropologi statunitensi, M. Mead e R. Birdwhistell (il padre della cinesica stessa), hanno dimostrato che solo alcune delle espressioni riprodotte dall’essere umano sono frutto del proprio retaggio culturale (ad esempio scuotere la testa per dissentire o fare il gesto dell’OK con la mano per approvare) [11]. Esistono infatti delle espressioni universali comuni a tutti gli esseri umani, indipendentemente da età, sesso o luogo di provenienza.

Vengono quindi identificate sei macro-categorie di espressioni comuni:

- Sorpresa,
- Gioia,
- Paura,
- Disgusto,
- Rabbia,

- Tristezza.

Due persone qualsiasi, in risposta a due stimoli analoghi, producono appunto la stessa espressione, che può essere ricercata in una delle categorie appena elencate.

Pur non essendo un'espressione vera e propria, a questa lista va aggiunta l'espressione neutra, mostrata in Figura 1, corrispondente ad una mancanza di una mimica facciale particolare.



Figura 1 Mimica facciale dell'espressione neutra [11]

### 1.3.1. Facial Action Coding System e Action Units

Il *Facial Action Coding System* (FACS) è un sistema per il riconoscimento di movimenti del volto comuni nell'essere umano, come ad esempio sorridere, alzare le sopracciglia o abbassare la mascella [12].

I movimenti vengono suddivisi in due grandi categorie:

- *Action Unit* (AU): insieme di azioni fondamentali per i quali sono stati identificati i muscoli coinvolti.
- *Action Descriptor* (AD): una serie di movimenti per i quali non sono ancora stati indicati con precisione i muscoli coinvolti.

Ad ogni AU viene assegnato un valore di intensità rappresentato con le lettere da A ad E, con A intensità minore (movimento appena accennato) ed E intensità maggiore. Attualmente risultano catalogate 46 AU che, insieme ad un relativo valore di intensità, definiscono un preciso metodo per riconoscere e misurare la tipologia e l'intensità di un'espressione. L'insieme delle AU è collezionato in un database di

riferimento (*Cohn-Kanade Database*), di cui si fornisce un breve estratto nella Figura 2 [13].















AU1  Inner brow raiser	AU2  Outer brow raiser	AU4  Brow Lowerer	AU5  Upper lid raiser	AU6  Cheek raiser
AU7  Lid tighten	AU9  Nose wrinkle	AU12  Lip corner puller	AU15  Lip corner depressor	AU17  Chin raiser
AU23  Lip tighten	AU24  Lip presser	AU25  Lips part	AU27  Mouth stretch	

Figura 2 Esempi di Action Unit presenti nel Cohn-Kanade Database [14]

## 1.4. Emozioni primarie

Successivi studi compiuti da Ekman affermano che non solo le espressioni sono comuni ad ogni essere umano, ma lo sono anche le emozioni che le causano, chiamate emozioni primarie ed in corrispondenza biunivoca con l'insieme di categorie identificato nel Paragrafo 1.3.

Per quanto riguarda l'espressione neutra non esiste però un'emozione collegata. L'essere umano prova costantemente emozioni e il volto assume una mimica facciale neutra nel caso in cui l'intensità dell'emozione provata non è tale da provocare un'espressione specifica sul volto [11].

Così come le espressioni che le esternano, le emozioni primarie sono inoltre riconosciute come un lascito evolutivo, piuttosto che un retaggio culturale, in quanto comuni al genere umano e riconoscibili anche in primati diversi dall'uomo [15].



#### **1.4.1. Sorpresa**

La sorpresa è la più breve delle emozioni primarie in termini di durata, riconosciuta come emozione transitoria. È caratterizzata, come mostrato in Figura 3, da occhi sgranati, mascella abbassata ed inarcamento verso l'alto delle sopracciglia, con la conseguente formazione di rughe orizzontali sulla fronte (più o meno evidenti a seconda dell'età).



Figura 3 Mimica facciale della sorpresa [11]

Si manifesta quando si viene sottoposti improvvisamente ad uno stimolo, nell'intervallo tra il momento della sottoposizione e la presa coscienza riguardo la natura dello stimolo stesso. In questo istante la sorpresa viene infatti soppiantata dall'emozione collegata allo stimolo iniziale.

Se però un'espressione di questo genere è mantenuta per un periodo di tempo prolungato ci si trova in presenza o di una finta sorpresa o di un'espressione di incredulità, non assimilabile con alcuna emozione.

#### **1.4.2. Gioia**

Con il termine gioia si indica un'emozione che scaturisce in seguito a diverse tipologie di stimoli.

Può infatti manifestarsi in seguito a sensazioni fisiche o mentali positive, interesse per qualcosa, cessazione di uno stimolo negativo o apprezzamento da e verso gli altri.

Nonostante la mimica corporale possa variare in risposta ai diversi tipi di stimoli, la mimica facciale mantiene sempre le stesse caratteristiche. È infatti riconoscibile dalla comparsa sul volto di un sorriso, come mostrato in Figura 4, o una risata, a seconda dell'intensità dell'emozione provata.



Figura 4 Mimica facciale della gioia [16]

#### **1.4.3. Paura, disgusto e rabbia**

La paura è quell'emozione che si prova quando si teme di subire un danno fisico o psicologico, il disgusto quella che si avverte in risposta ad un sentimento di repulsione, mentre la rabbia è collegata a sentimenti di odio o sdegno verso qualcuno o qualcosa.

Nella trattazione queste tre emozioni sono raggruppate in quanto, come anche mostrato nelle Figure 5, 6 e 7, la mimica facciale riprodotta è molto simile nei tre casi. In tutte l'elemento predominante è l'abbassamento e l'avvicinamento tra loro delle sopracciglia. Nelle prime due abbiamo un sollevamento del labbro superiore, anche solo accennato, che rimane però serrato nella rabbia (Figura 5). È proprio del disgusto l'arricciamento del naso (Figura 6), mentre la paura è caratterizzata da occhi sgranati e palpebre innalzate (Figura 7).

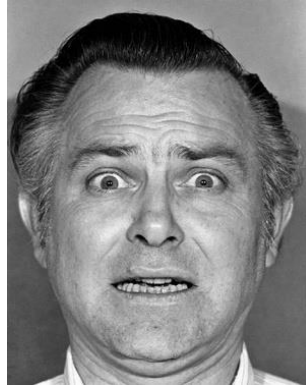


Figura 5 Mimica facciale della paura [16]

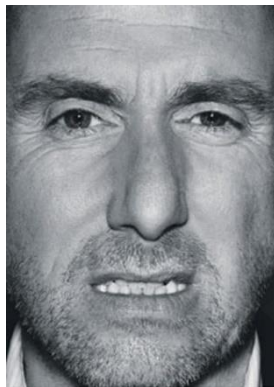


Figura 6 Mimica facciale del disgusto [17]



Figura 7 Mimica facciale della rabbia [16]

#### **1.4.4. Tristezza**

La tristezza è l'emozione associata alla malinconia e dolore psicologico. È l'emozione più duratura e si protrae anche una volta cessata la relativa mimica facciale. Per questo motivo risulta complicato rilevare la tristezza attraverso le

espressioni del volto, in quanto non sempre l'emozione è associata ad un'esternazione di questo tipo.

Nei primi momenti di manifestazione dell'emozione è comunque possibile osservare variazioni nella mimica facciale. Come mostrato in Figura 8, è riconoscibile dall'abbassarsi delle sopracciglia e dello sguardo in generale. Per quanto concerne la parte bassa del volto non si ha una discriminante precisa poiché l'espressione della bocca in alcuni casi può ricordare il disgusto, mentre in altri può rimanere inespressiva.



Figura 8 Mimica facciale della tristezza [18]

## 1.5. Scelta delle emozioni rilevate

L'applicazione sviluppata rileva quindi la sorpresa, la gioia e il disgusto, oltre all'espressione neutra, usata come punto di riferimento iniziale per il riconoscimento delle altre emozioni.

Dall'insieme delle emozioni rilevate è stata esclusa la tristezza in quanto per sua natura emozione difficile da rilevare attraverso la mimica facciale.

In relazione alla precisione delle API fornite con l'hardware selezionato per lo sviluppo (maggiori dettagli a riguardo nel capitolo successivo), si è scelto di rilevare una sola emozione tra rabbia, paura e disgusto. Concentrandosi sulle sopracciglia si capisce che l'emozione rilevata è una di queste tre, ma non si ha una precisione tale per eseguire un ulteriore campionamento basandosi sulle minime differenze che

queste emozioni presentano. In definitiva, tra le tre, si rileva il disgusto in quanto emozione più facile da mimare o provocare in uno scenario di test.

## 2 Kinect

Il sensore Kinect (nome in codice “Project Natal”) [19] è un dispositivo sviluppato da Microsoft Corporation e rilasciato nella sua prima versione nel novembre 2010, pensato originariamente per integrarsi con la console Xbox 360, con la finalità di poter interagire con la console stessa tramite il movimento del corpo, senza bisogno di impugnare altri controller.

Il sensore è dunque in grado di identificare l’utente e riconoscerne i movimenti del corpo nello spazio circostante.

Al momento della trattazione sono state rilasciate quattro versioni dell’hardware:

- *Kinect*, la prima versione, rilasciata per Xbox 360;
- *Kinect For Windows*, la prima versione puramente per sviluppatori;
- *Kinect One*, versione rilasciata per piattaforma Xbox One;
- *Kinect For Windows 2.0*, seconda versione per sviluppatori.

### 2.1. Componenti hardware

Il sensore Kinect è composto, come mostrato nella Figura 9 da [20]:

1. un proiettore laser infrarossi e un sensore CMOS, che combinati permettono sia la rilevazione tridimensionale dell’ambiente, sia la rilevazione dei movimenti dell’utente;
2. una telecamera RGB per l’acquisizione del video;
3. un array di microfoni utilizzati per il riconoscimento vocale;
4. un supporto motorizzato per l’inclinazione automatica del dispositivo.

La prima versione rilasciata del dispositivo è in grado di identificare un massimo di 4 persone in uno spazio compreso tra gli 0.8 e i 4 metri dall’obiettivo, mentre l’array di microfoni riesce ad escludere fino a 20dB di rumore ambientale, 26dB se il suono proviene da dietro l’apparecchio.

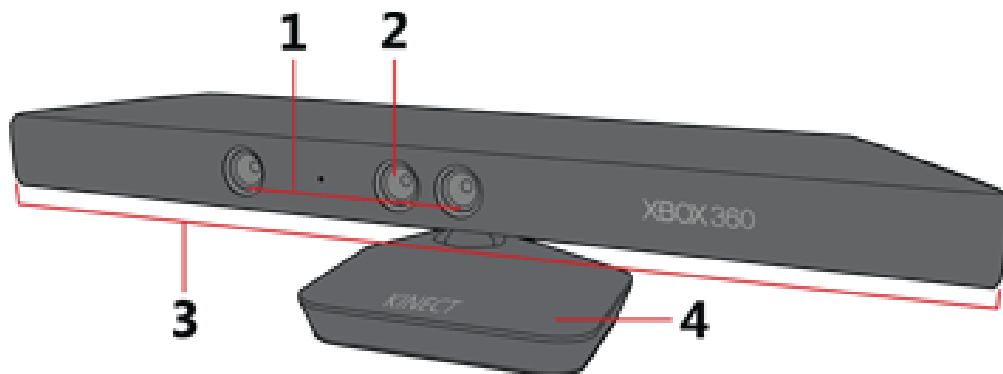


Figura 9 Componenti del sensore Kinect

## 2.2. Tecniche per l'acquisizione di immagini 3D

Allo stato dell'arte esistono diverse tecniche per ricavare un'immagine 3D, le più utilizzate sono [21]:

- *triangolazione stereo*, calcola la profondità degli oggetti a partire da due immagini riprese da telecamere differenti;
- *tecnica del time-of-flight*, grazie alla quale è possibile ottenere la profondità quantificando il tempo che un fascio di luce emesso impiega per tornare indietro, in seguito alla collisione con un oggetto;
- *emissione di pattern luminosi*, tecnica che consiste nel proiettare un pattern luminoso noto e, basandosi sulla distorsione di quest'ultimo sugli oggetti, è possibile risalire alla struttura tridimensionale dell'immagine.

Anche per quanto riguarda l'emissione di pattern luminosi si può fare un'ulteriore suddivisione. Si possono infatti proiettare [21]:

- *linee*, esempio nella Figura 10, soluzione non molto efficiente e soggetta ad errori su soggetti in movimento;
- *pattern periodici*, esempio nella Figura 11, soluzione più efficiente della precedente, ma non in grado di lavorare su lunghe distanze a causa della distorsione subita del pattern stesso;

- *pattern pseudo-casuali*, esempio nella Figura 12, soluzione, brevettata nel 2005 da Zalevsky, Shpunt, Maizels e Garcia [22], che permette di ottenere risultati in maniera semplice ed efficiente, senza incorrere negli errori della precedente soluzione grazie alla pseudo-casualità del pattern stesso.



Figura 10 Esempio di proiezione di linee

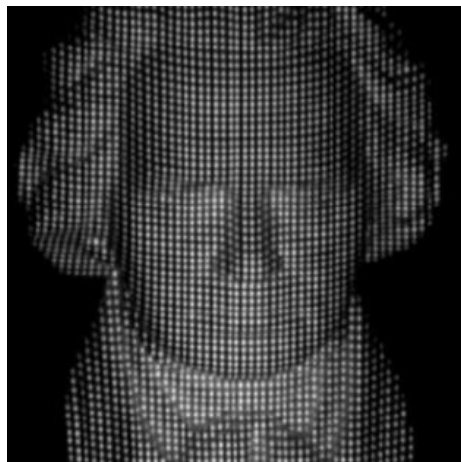


Figura 11 Esempio di proiezione di un pattern periodico





Figura 12 Esempio di proiezione di un pattern pseudo-casuale [19]

Per l'ultima versione rilasciata della Kinect in ambito videoludico (Kinect One al momento della trattazione) Microsoft adotta una soluzione time-of-flight, sviluppata dalla compagnia statunitense Canesta [23].

Per la Kinect For Windows (utilizzata ai fini del progetto) è utilizzata l'ultima tipologia di pattern descritta (pattern pseudo-casuali) e il relativo sensore, realizzato dalla compagnia israeliana PrimeSense. Viene proiettato in questo caso un pattern luminoso infrarosso che successivamente è rilevato dal sensore CMOS sincronizzato sulla stessa frequenza. È quindi necessario il solo calcolo dello scostamento tra i punti ripresi e quelli proiettati in origine per poter costruire l'immagine 3D [24].

Questo tipo di tecnologia permette di suddividere l'area ripresa dalla telecamera in tre regioni differenti, come mostrato in Figura 13: una prima da 0.8 a 1.2 metri, una seconda da 1.2 a 3.5 metri ed una terza da 3.5 a 4 metri, con la seconda regione dotata di maggior precisione.

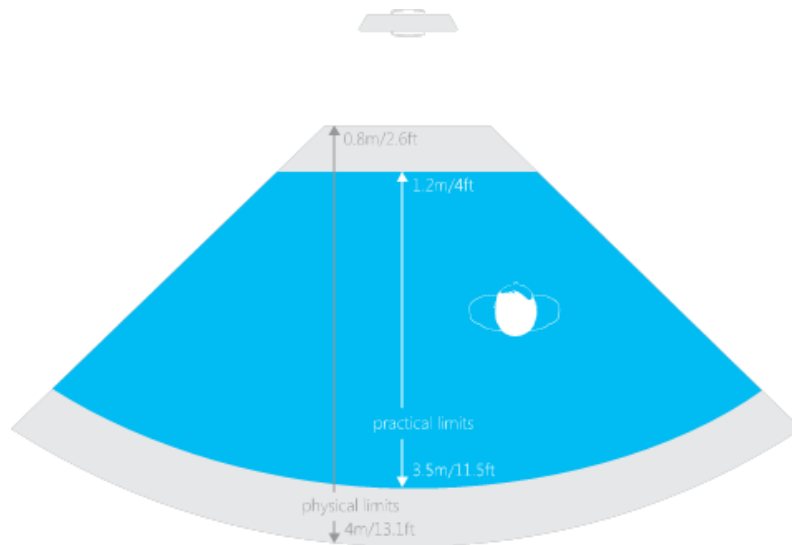


Figura 13 Raggio d'azione del sensore Kinect For Windows [25]

### 2.3. Kinect For Windows

Nel passare degli anni l'avanzamento tecnologico e il maggiore interesse per il prodotto, anche all'infuori del campo videoludico, hanno portato Microsoft a rilasciarne nuove versioni.

Nel febbraio 2012 vede la luce Kinect For Windows ovvero una versione dell'hardware pensata per essere interfacciata con un computer e non con una console.

Il prodotto presenta le stesse componenti hardware del suo predecessore, ma la vera differenza è contenuta nel SDK rilasciato in concomitanza col prodotto stesso. Questo aggiunge migliorie nel tracciamento del corpo e della voce oltre ad una *near mode* che permette una buona acquisizione dell'immagine ad una distanza non inferiore ai 40 cm, migliorando così di 40 cm il limite minimo per una buona rilevazione individuato dal prodotto precedente. In questo caso la regione con maggior precisione si colloca tra i 0.8 ed i 2.5 metri dall'obiettivo, come mostrato in Figura 14.

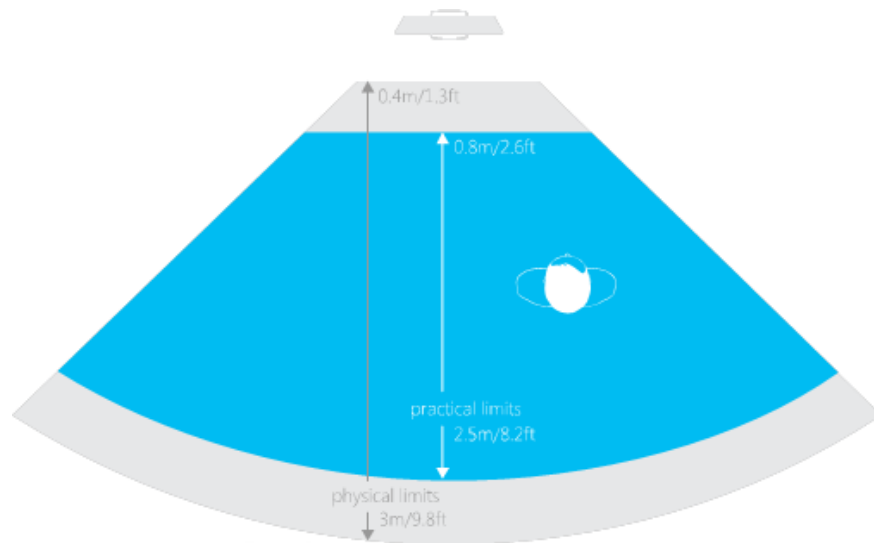


Figura 14 Raggio d'azione del sensore Kinect For Windows con *near mode* [25]

Quest'ultima funzionalità ha portato alla scelta di Kinect For Windows per lo svolgimento del progetto, piuttosto che altre versioni più recenti, dotate di un hardware migliore ma sprovviste di *near mode*, funzionalità di interesse per il progetto stesso [23].

## 2.4. SDK utilizzati

Nei seguenti sottoparagrafi vengono descritti gli SDK utilizzati per lo sviluppo del progetto.

### 2.4.1. Kinect For Windows SDK 1.8

Per lo sviluppo dell'applicazione si è scelto di usare il Kinect For Windows SDK 1.8 e relativo Developer Toolkit [26], ovvero la versione più recente pensata per il modello di Kinect usato nello sviluppo del progetto.

Con l'utilizzo di questo specifico SDK bisogna far attenzione ad errori ed imprecisioni dovuti a limitazioni tecnologiche e non ancora risolti al momento della trattazione.

Questo richiede che il soggetto osservato venga posizionato il più possibile al centro dell'inquadratura, onde evitare imprecisioni nella rilevazione, tanto più accentuate

quanto più il soggetto stesso si discosta dal centro. Inoltre va evitata la presenza di fonti luminose nel raggio di azione della Kinect, che potrebbe comportare la perdita del tracciamento del soggetto e quindi una perdita di dati.

#### 2.4.2. Face Tracking SDK

Il Face Tracking SDK [27], in concomitanza con il Kinect For Windows SDK, permette di dedurre la posizione della testa insieme ad altre informazioni di base relative al volto umano e le rende fruibili al programmatore tramite delle API.

I dati principali restituiti sono quindi di 3 tipi [28]:

- *Head Pose Angles*: una terna di angoli compresi tra  $-90^\circ$  e  $90^\circ$ , rappresentanti l'inclinazione della testa rispetto ad una posizione prestabilita in cui si guarda fissi nell'obiettivo, come mostrato in Figura 15;
- *Tracked Points*: un array di 100 punti che identificano gli elementi principali che caratterizzano il volto, distribuiti come mostrato in Figura 16;
- *Animation Units*: di particolare interesse per lo sviluppo del programma e descritti in dettaglio nel paragrafo successivo.

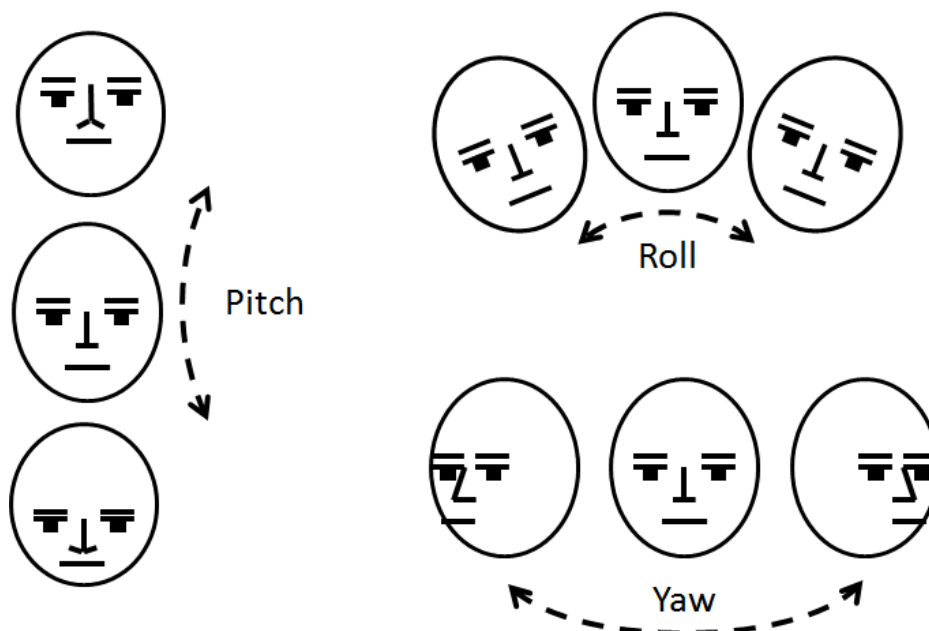


Figura 15 Movimenti descritti dagli *Head Pose Angles* [28]

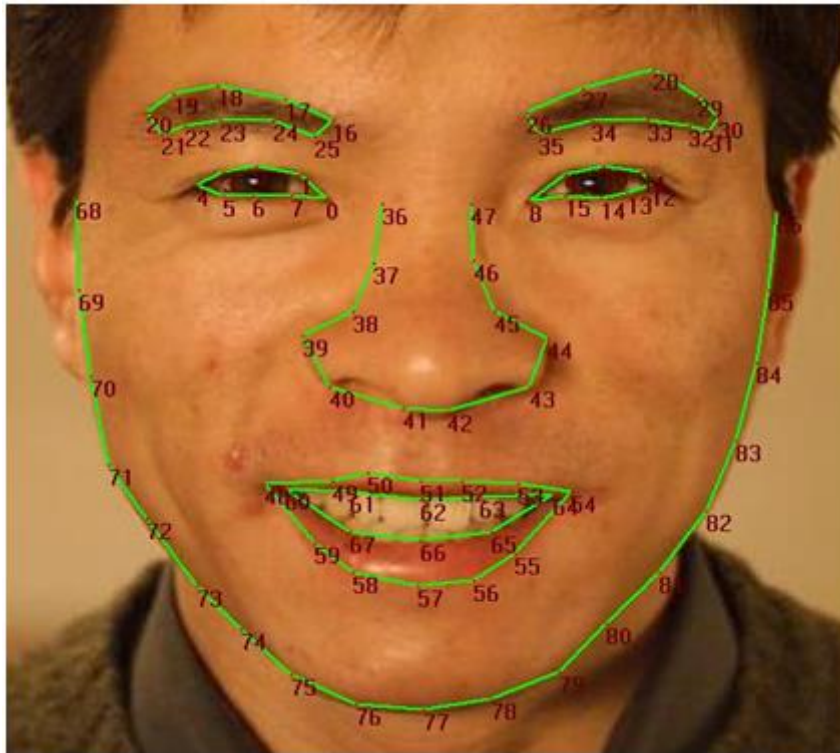


Figura 16 Posizione di alcuni dei 100 *Tracked Points* [28]

#### 2.4.2.1. Animation Units

Le *Animation Unit* (AU) [28] sono un sottoinsieme delle *Action Unit* più comuni. Questo sottoinsieme è espresso sotto forma di un array restituito dal metodo `EnumIndexableCollection<AnimationUnit,float> GetAnimationUnitCoefficients()` contenuto nella libreria `Microsoft.Kinect.Toolkit.FaceTracking.FaceTrackFrame`.

A differenza delle *Action Unit*, i valori delle *Animation Unit* non sono rappresentati da lettere dell'alfabeto, ma da un numero reale compreso tra -1 e 1.

Nella Tabella 1, riportata di seguito, vengono descritte le *Action Unit* contenute nell'array delle *Animation Unit* restituito, con la relativa posizione nello stesso e il significato dei valori assegnati.



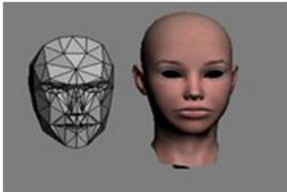


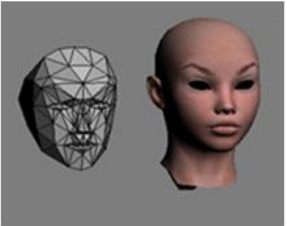
Nome dell'AU	Posizione nell'array	Avatar	Valori e interpretazione
Upper Lip Raiser	0		0 = neutro, denti coperti 1 = mostrare i denti completamente -1 = labro posizionato il più giù possibile
Jaw Lowerer	1		0 = bocca chiusa 1 = bocca completamente aperta -1 = bocca chiusa (equivalente a 0)
Lip Stretcher	2		0 = neutro 1 = sorriso accentuato al massimo -1 = labbra che mimano un bacio
Brow Lowerer	3		0 = neutro 1 = sopracciglia completamente abbassate -1 = sopracciglia completamente alzate
Lip Corner Depressor	4		0 = neutro 1 = posizione della bocca triste -1 = posizione della bocca sorridente
Outer Brow Raiser	5		0 = neutro 1 = sopracciglia completamente alzate -1 = sopracciglia completamente abbassate

Tabella 1 Classificazione delle Animation Unit [26]

Le *Animation Unit* scelte per il riconoscimento delle emozioni di interesse al progetto vengono analizzate in dettaglio nel capitolo relativo all'implementazione.

## **2.5. Perché si è scelta la Kinect**

La Kinect offre quindi una serie di importanti funzionalità per quanto riguarda la rilevazione di immagini 3D e dei movimenti del corpo nella sua totalità e non solo del volto, che però non vengono usati nello svolgimento del progetto, poiché non necessari allo scopo.

Lo stesso si sarebbe potuto raggiungere anche con l'utilizzo di un hardware differente, senza bisogno che questo debba essere dotato di funzionalità 3D.

Si è scelto di utilizzare comunque il sensore Kinect dato che offre una vasta gamma di librerie specifiche per il riconoscimento facciale in continua evoluzione, costante aggiornamento ed usate da un vasto numero di sviluppatori in tutto il mondo [29]. Inoltre lo scopo prefissato non è di sviluppare nuove librerie, ma quello di implementare un programma che, partendo da API consolidate, stabili e performanti, riconosca l'emozione di una persona.

Un altro fattore che ha portato alla scelta di questo dispositivo è la sempre maggiore diffusione che quest'ultimo sta avendo, in quanto accessorio di una delle console videoludiche di maggior fama e diffusione al momento della trattazione. La Kinect è stata infatti premiata con il MacRobert Award nel 2011 [30] e, stando alle statistiche rilasciate dalla casa produttrice, il numero di unità vendute nel febbraio 2013 si attestava addirittura sui 24 milioni [31].

### 3 Implementazione

In questo capitolo vengono descritte le scelte compiute in fase di sviluppo dell'applicazione per il riconoscimento delle emozioni, insieme ai metodi principali dell'applicazione stessa.

L'applicazione è stata scritta in linguaggio C# usando le librerie fornite da Microsoft descritte nel Capitolo 2 ed è pensata per funzionare su sistemi operativi Windows dotati dei driver per l'utilizzo della Kinect.

Come IDE si è scelto di utilizzare *Microsoft Visual Studio* in quanto ambiente di sviluppo di riferimento per il linguaggio scelto.

Per la parte grafica si fa uso del framework *Windows Presentation Foundation*, descritto in dettaglio nel paragrafo successivo.

#### 3.1. Windows Presentation Foundation

*Windows Presentation Foundation* (WPF) [32] è un insieme di librerie contenute nel *Framework .NET* a partire dalla versione 3.0, usate per lo sviluppo della parte relativa all'interfaccia grafica dell'applicazione. Queste librerie permettono di creare interfacce di ogni genere in ambiente Windows, anche all'interno di browser. WPF fa inoltre riferimento alle librerie *DirectX*, sfruttando quindi l'accelerazione hardware fornita dalle moderne schede grafiche.

Le interfacce sono programmate usando un linguaggio specifico, basato su XML denominato XAML (*eXtensible Application Markup Language*) [33].

Nello specifico dell'applicazione sviluppata è stata usata la versione di WPF contenuta nel *Framework .NET 4.0*.



### 3.2. Panoramica dell'applicazione

Nelle Tabelle 2 e 3 vengono elencate le classi e le pagine XAML di cui si compone l'applicazione, accompagnate da una breve descrizione del loro scopo e funzionamento.

Classe	Descrizione
EmotionsAnalyzer.cs	Classe predisposta al riconoscimento delle emozioni
FaceTrackingViewer.cs	Classe delegata alla visualizzazione del video ricevuto dalla Kinect e al disegno dei punti tracciati sul volto dell'utente
SkeletonFaceracker.cs	Classe che estrae le informazioni trasmesse dalla Kinect
App.xaml.cs	Classe contenente il codice relativo alla pagina XAML con lo stesso nome.
MainWindow.xaml.cs	
MainPage.xaml.cs	
InstructionPage.xaml.cs	
PresentationPage.xaml.cs	

Tabella 2 Elenco delle classi dell'applicazione

Pagina XAML	Descrizione
App.xaml	Pagina di partenza di un'applicazione WPF [33]
MainWindow.xaml	Contenitore delle varie pagine dell'applicazione
MainPage.xaml	Pagina principale
InstructionPage.xaml	Pagina contenente le istruzioni dell'applicazione
PresentationPage.xaml	Pagina iniziale contenente la presentazione dell'applicazione

Tabella 3 Elenco delle pagine XAML dell'applicazione

Ai fini della trattazione vengono approfondite esclusivamente la pagina *MainPage.xaml* e la classe *EmotionsAnalyzer.cs*, poiché ritenute di maggiore interesse. Viene tralasciato, ad esempio, il codice relativo all'utilizzo del sensore Kinect, facilmente reperibile nella documentazione fornita da Microsoft.

### 3.3. Aggiunta del metodo *GetShapePoints*

Prima di iniziare la trattazione del codice vero e proprio dell'applicazione, si segnala un cambiamento apportato in fase di tesi alle API utilizzate, tramite l'aggiunta del metodo *public void GetShapePoints(ref Vector2DF[] vector2DF)*, riportato nella Tabella 4, alla libreria *Microsoft.Kinect.Toolkit.FaceTracking.FaceTrackFrame* [35].

```
public void GetShapePoints(ref Vector2DF[] vector2DF)
{
    IntPtr ptBuffer = IntPtr.Zero;
    uint ptCount = 0;
    this.faceTrackingResultPtr.Get2DShapePoints(out ptBuffer, out ptCount);
    if (ptCount == 0)
    {
        vector2DF = null;
        return;
    }
    if (vector2DF == null || (vector2DF != null && vector2DF.Length !=
                                                                    ptCount))
    {
        vector2DF = new Vector2DF[ptCount];
    }
    ulong sizeInBytes = (ulong)Marshal.SizeOf(typeof(Vector2DF));
    for (ulong i = 0; i < ptCount; i++)
    {
        vector2DF[i] = (Vector2DF)Marshal.PtrToStructure((IntPtr)((ulong)ptBuffer
                                                                    + (i * sizeInBytes)),typeof(Vector2DF));
    }
}
```

Tabella 4 Metodo GetShapePoints

Il motivo di questa modifica è da ricercarsi in un'impresione contenuta nella documentazione fornita da Microsoft. L'array *Tracked Points* (in dettaglio nel Paragrafo 2.4.2) viene infatti detto contenere i 100 punti descritti in Figura 16. In realtà questo array contiene un insieme di 121 punti che identificano i vertici della griglia del modello Candide-3, maschera usata per il riconoscimento di parametri standard sul volto umano [36].

Il metodo *GetShapePoints* restituisce i 100 punti così come indicato nella documentazione. La riproduzione di singoli punti sul volto, invece di una griglia, produce inoltre un effetto più chiaro all'utente, permettendo di identificare l'esatta posizione dei punti e delle loro corrispondenze sul proprio volto, non coprendone l'immagine complessiva.

La differenza appena descritta può essere apprezzata paragonando la Figura 17 con la Figura 18.

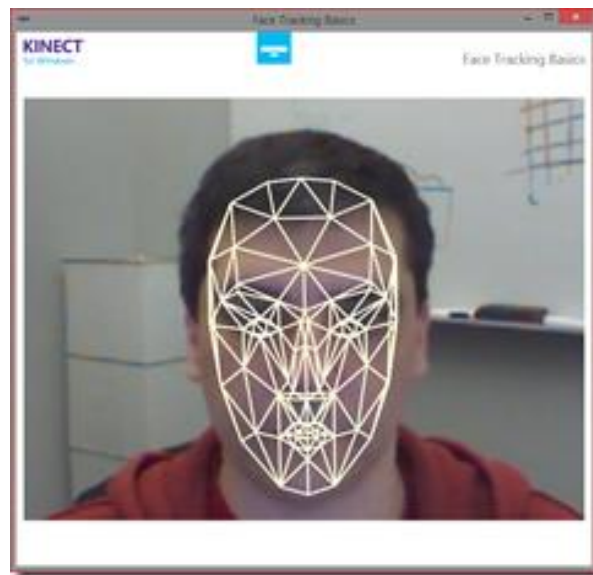


Figura 17 Esempio di proiezione del modello Candide-3 [35]

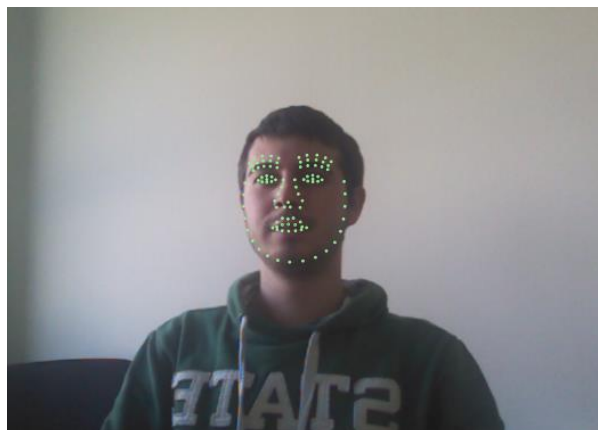


Figura 18 Esempio di proiezione dei *Tracked Points*

### 3.4. Interfaccia grafica

Nella Figura 19 è mostrata la pagina principale dell'applicazione, sviluppata con l'utilizzo di WPF, e sono individuate le tre sezioni che la compongono:

- A. Video e punti tracciati;
- B. *TextBox*;
- C. Controlli.

Nella sezione A vengono riprodotti i punti del volto tracciati dalla Kinect, sovrapposti al contenuto ripreso dalla telecamera RGB del sensore stesso. Questi sono del tutto indipendenti tra loro e la sovrapposizione è pensata per far coincidere i punti rilevati con i relativi punti sul volto della persona ripresa.

Per quanto riguarda i punti del volto viene richiamata la classe *FaceTrackingViewer*, che a sua volta invoca la classe *SkeletonFaceTracker* per rilevare i suddetti punti e poi disegnarli.

Nella sezione B è presente una *TextBox*, in cui è stampato a video l'output e altre informazioni che il programma restituisce all'utente, come l'inizio o la fine del setup dell'applicazione.

Nella sezione C sono presenti i comandi dell'applicazione. I due *RadioButton*, sono usati per mettere in pausa o far partire la rilevazione, mentre i due bottoni presenti permettono di cancellare completamente il contenuto della *TextBox* ("RESET") e di salvare il contenuto della stessa in un file di testo ("SAVE").

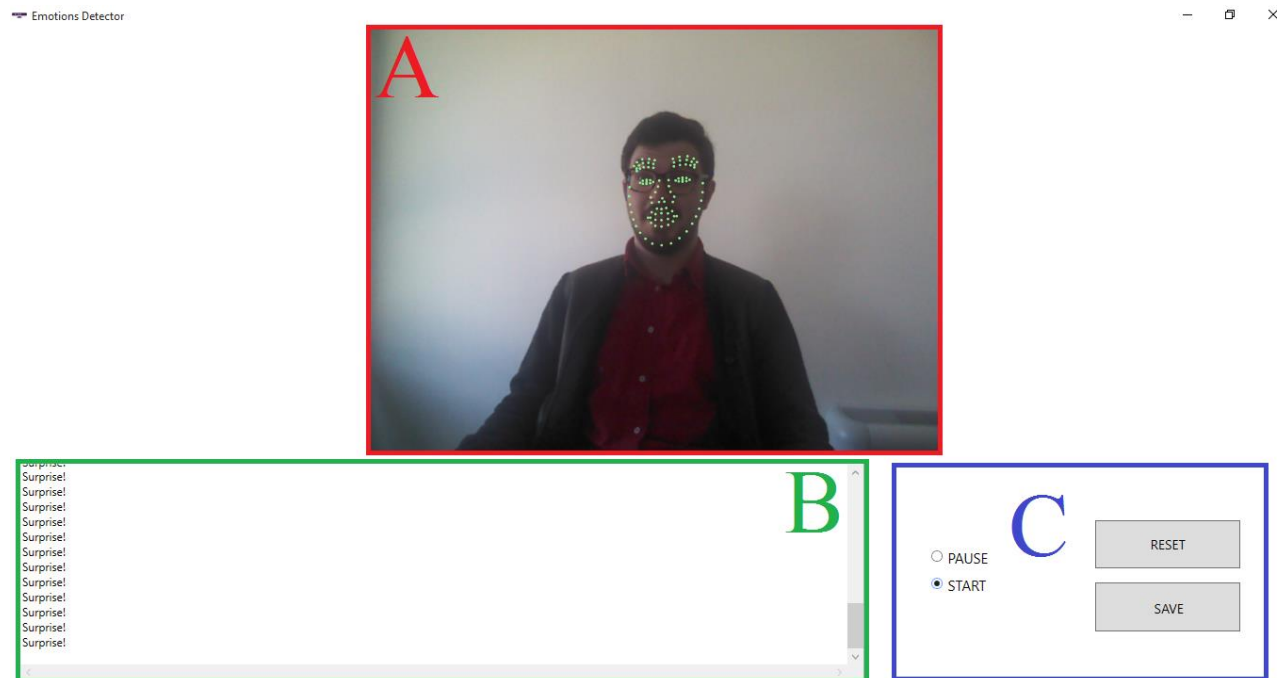


Figura 19 Schermata principale dell'applicazione

### 3.5. EmotionsAnalyzer

La classe *EmotionsAnalyzer* è considerata la più importante all'interno dell'applicazione, occupandosi di processare le informazioni riguardanti il volto ricevute dalla Kinect, così da poter rilevare l'emozione.

Ai fini dell'implementazione sono state usate le librerie elencate nella Tabella 5 con le relative funzionalità.

Libreria	Descrizione
Microsoft.Kinect.Toolkit.FaceTracking	Libreria per il tracciamento del volto
System	Libreria per l'utilizzo di stringhe
System.Collections.Generic	Libreria per l'utilizzo di collezioni

Tabella 5 Librerie utilizzate nella classe *EmotionsAnalyzer*

Prima del riconoscimento vero e proprio dell'emozione la classe esegue un setup iniziale, calcolando dei parametri di base sui cui poi basare la rilevazione (maggiori dettagli nei paragrafi successivi). Questi valori sono memorizzati in una serie di variabili d'istanza, elencate nelle Tabella 6 e Tabella 7. Le liste indicate nella prima sono utilizzate per raccogliere una serie di campioni, sulla base dei quali vengono calcolati i valori iniziali, salvati poi nelle variabili indicate in Tabella 7, per l'identificazione dell'emozione.

```
private List<float> smileSetupArray;
private List<float> browUpSetupArray;
private List<double> browLowRSetupArray;
private List<double> browLowLSetupArray;
```

Tabella 6 Liste usate per il setup dell'applicazione

```
private float smileBasicValue;
private float browUpBasicValue;
private double browLowRBasicValue;
private double browLowLBasicValue;
```

Tabella 7 Variabili rappresentanti i valori base per la rilevazione

Nella Tabella 8 vengono invece riportati i metodi principali della classe con una breve descrizione delle loro funzionalità principali, due dei quali verranno ripresi più avanti nella trattazione e descritti nel dettaglio.

Metodo	Descrizione
<code>internal String Setup(FaceTrackFrame frame)</code>	Si occupa di eseguire il setup iniziale dell'applicazione
<code>private void EvaluateInitialValues()</code>	Calcola i valori di base per la rilevazione
<code>private double PointDistance(PointF pf1 , PointF pf2)</code>	Calcola la distanza tra due oggetti PointF
<code>internal String AnalyzeEmotion(FaceTrackFrame frame)</code>	Esegue la rilevazione dell'emozione

Tabella 8 Metodi principali della classe *EmotionsAnalyzer*

### **3.5.1. Rilevazione attraverso Animation Units**

Come descritto nel Capitolo 2, si utilizzano alcune AU (*Animation Units*) per identificare le emozioni. Queste verranno analizzate nello specifico, indicando quali si sono utilizzate ai fini dello sviluppo e quali no.

Per la lista completa delle AU fornite dalle librerie di utilizzare si rimanda alla Tabella 1.

#### **3.5.1.1. Test eseguiti sulle Animation Units**

Con lo scopo di verificarne la precisione, durante la fase di implementazione sono stati svolti dei test sui valori contenuti nell'array delle AU, che come già accennato nel Capitolo 2 sono numeri reali compresi tra -1 ed 1.

Si è dunque sviluppato un semplice programma con il solo scopo di stampare a video questi valori. Sono stati selezionati 10 soggetti diversi, ai quali è stato chiesto di mimare le espressioni descritte dalle AU, partendo da un'espressione neutra e con intensità crescente, così da poter misurare l'andamento dei valori restituiti in risposta alle diverse mimiche facciali.

I test svolti in questo modo hanno dimostrato che i valori restituiti in risposta alla stessa mimica facciale, discostavano molto tra soggetto e soggetto. Ad esempio un sorriso, AU con posizione 2, veniva associato al valore 0.4 in alcuni soggetti, mentre in altri raggiungeva il valore 1.

Questi risultati hanno però evidenziato che le variazioni tra l'espressione neutra e l'espressione desiderata erano all'incirca le stesse.

Per questo motivo si è quindi preferito basarsi su queste variazioni, piuttosto che sui valori assoluti delle AU.

Nella Tabella 9 e Tabella 10 sono riportate le rilevazioni fatte rispettivamente per l'AU *Lip Stretcher* e l'AU *Outer Brow Raiser*.

Soggetto	1	2	3	4	5	6	7	8	9	10
<b>Valore Iniziale</b>	0,1	0,4	0,3	0,1	-0,3	-0,3	0,5	0,2	0,4	-0,2
<b>Valore Finale</b>	0,6	0,9	0,8	1	0,2	-0,1	0,9	0,9	1	0,2
<b>Variazione</b>	0,5	0,5	0,5	0,9	0,5	0,2	0,4	0,7	0,6	0,4

Tabella 9 Rilevazioni eseguite per l'AU *Lip Stretcher*

Soggetto	1	2	3	4	5	6	7	8	9	10
<b>Valore Iniziale</b>	0	0,3	-0,5	0,1	-0,1	-0,2	-0,1	-0,4	-0,1	0,1
<b>Valore Finale</b>	0,4	0,5	-0,4	0,2	0,1	0,1	0,1	-0,2	0,3	0,2
<b>Variazione</b>	0,4	0,2	0,1	0,1	0,2	0,3	0,2	0,2	0,4	0,1

Tabella 10 Rilevazioni eseguite per l'AU *Outer Brow Raiser*

### 3.5.1.2. Margini di riferimento

La media tra le variazioni ottenute con i test descritti precedentemente è utilizzata come valore discriminante per l'identificazione di un'emozione. Questi valori, descritti in dettaglio nella Tabella 11, vengono definiti *Margini di Riferimento*.

Variabili	Emozione di riferimento
private static readonly float SMILE_GAP = 0.5F;	Gioia
private static readonly float BROW_UP_GAP = 0.2F;	Sorpresa
private static readonly double BROW_LOW_GAP = 150;	Disgusto

Tabella 11 Margini di riferimento

La differenza tra il tipo dei valori relativi al disgusto e il tipo dei valori relativi a gioia e sorpresa è dovuta all'uso di una tecnica di rilevazione del disgusto che non si basa sulle informazioni contenute nelle AU e che viene descritta in seguito.



### 3.5.1.3. Animation Units utilizzate

Non sono utilizzati per il rilevamento delle emozioni le AU con posizione nell'array 0 ed 1, rispettivamente *Upper Lip Raiser* e *Jaw Lowerer*, poiché la prima è risultato essere troppo poco sensibile, mentre la seconda, al contrario, troppo sensibile. Inizialmente si è pensato di usare l'AU *Jaw Lowerer* per il riconoscimento della sorpresa, ma il suo valore raggiunge il massimo anche quando si ride con la bocca aperta, rischiando di creare quindi confusione tra sorpresa e gioia.

Viene scartata anche l'AU con posizione 4, *Lip Corner Depressor*, poiché rileva un'espressione facciale innaturale e quindi di poca utilità.

Merita approfondimenti l'AU con posizione 3, *Brow Lowerer*, con la quale si pensava in principio di rilevare il disgusto. I valori di questa AU sembrano a prima vista essere esattamente opposti a quelli contenuti nell'AU con posizione 5, *Outer Brow Raiser*, almeno secondo la documentazione di riferimento, condizione poi smentita dai test effettuati. Il valore dell'AU *Brow Lowerer* in due espressioni "neutre" (sopracciglia rilassate) consecutive, intervallate da una in cui si sono abbassate le sopracciglia, discosta di molto. Vista quindi l'impossibilità di potersi assestare su un valore di riferimento per l'espressione neutra sul quale eseguire le rilevazioni, si è deciso di riconoscere il disgusto tramite l'esame di altri valori.

In definitiva si usano le AU con posizione 2 e 5 (*Lip Stretcher* e *Outer Brow Raiser*) per rilevare rispettivamente la gioia e la sorpresa.

### 3.5.2. Scelta per la rilevazione del disgusto

Appurata l'impossibilità di una rilevazione accurata del disgusto basandosi sui valori contenuti nelle AU, lo si rileva attraverso la misurazione della distanza tra due punti sulle sopracciglia e un terzo punto fisso.

Facendo riferimento ai punti mostrati in Figura 20 vengono presi come riferimento per le sopracciglia i punti 25 e 35, posizionati nella parte interna delle stessa, nell'angolo inferiore. Come terzo punto di riferimento fisso si usa il punto 89, posizionato sulla punta del naso, in quanto posizionato in una parte del volto che non prende parte nella mimica di emozioni diverse dal disgusto stesso, che è appunto caratterizzato tra le altre cose dall'arricciamento del naso.

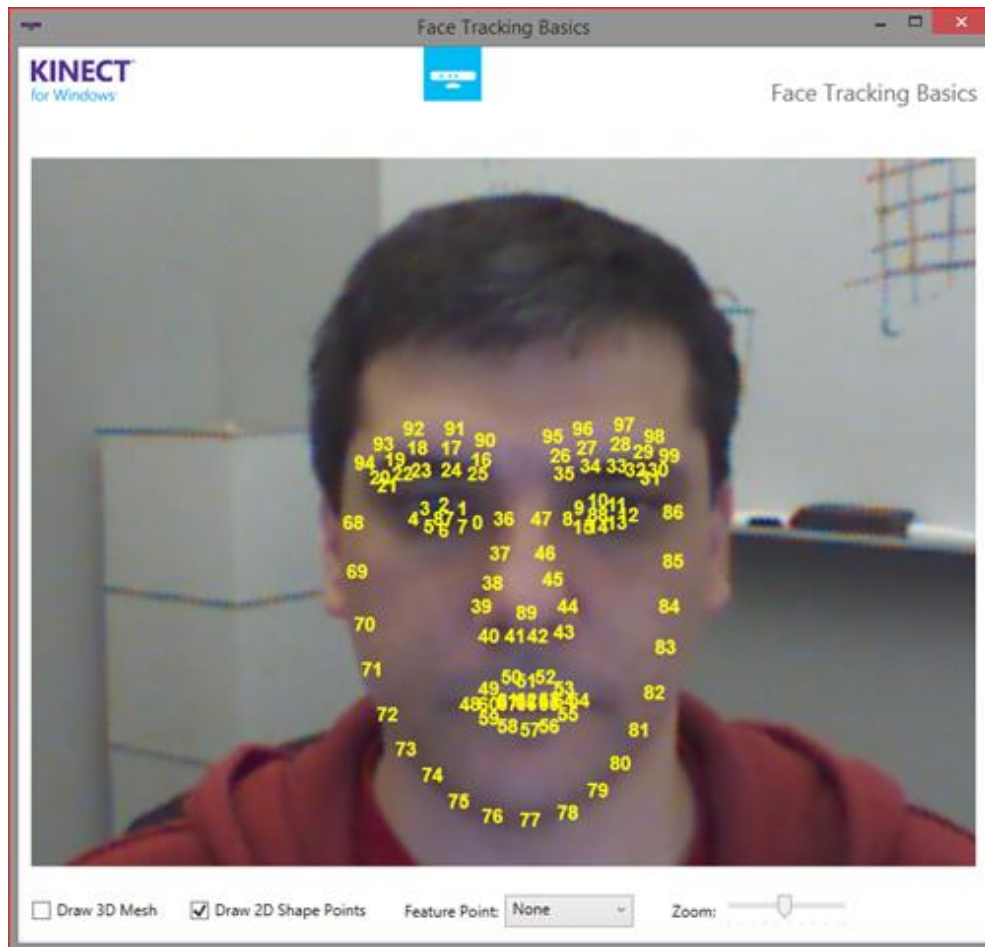


Figura 20 10 Elenco completo dei *Tracked Points* con relativo numero [35]

Facendo uso del metodo *private double PointDistance(PointF pf1 , PointF pf2)* tra i punti 25 e 35 con il punto 89 si calcolano durante il setup i valori di base per queste distanze. Lo stesso metodo viene inoltre utilizzato per calcolare le stesse distanze tra i punti nella fase decisionale riguardo l'emozione rilevata.

Necessitando di due distanze differenti è indicato con *BrowLowR* tutto ciò che riguarda l'abbassamento della sopracciglia destra e con *BrowLowL* tutto ciò che riguarda l'abbassamento della sinistra.

### 3.5.3. Setup dell'applicazione

Durante il setup dell'applicazione viene richiamato il metodo *internal String Setup(FaceTrackFrame frame)*, riportato nella Tabella 12.

```

internal String Setup(FaceTrackFrame frame)
{
    if (this.smileSetupArray.Count < SETUP_VALUES)
    {
        this.smileSetupArray.Add(frame.GetAnimationUnitCoefficients()[2]);
        this.browUpSetupArray.Add(frame.GetAnimationUnitCoefficients()[5]);
        this.browLowRSetupArray.Add(this.pointDistance(frame.GetShapePoints()[89],
                                                         frame.GetShapePoints()[25]));
        this.browLowLSetupArray.Add(this.pointDistance(frame.GetShapePoints()[89],
                                                         frame.GetShapePoints()[35]));

        return (".");
    }
    else
    {
        this.EvaluateInitialValues();
        this.isSetupComplete = true;
        return ("\r\nSetup Complete!\r\n");
    }
}

```

Tabella 12 Metodo Setup

Il metodo riceve in input un oggetto *FaceTrackFrame*, contenente le informazioni riguardanti il volto tracciato. Ha il compito di rilevare tanti campioni quanti indicati nella variabile *SETUP\_VALUES*, nell'applicazione sviluppata si rilevano i primi 50 campioni, e di salvarli nelle liste descritte nella Tabella 6. Una volta collezionati tutti i campioni richiesti viene richiamato il metodo *EvaluateInitialValues()*, con lo scopo di calcolare la mediana sui valori richiesti. I risultati saranno proprio i valori base di partenza per la rilevazione delle emozioni e come tali verranno salvati nelle variabili d'istanza descritte nella Tabella 7.

Per dare all'utente un senso di avanzamento durante il setup il metodo restituisce in questa fase un singolo punto, che viene stampato nella *TextBox* nella schermata principale dell'applicazione. Ottenendo quindi un punto per ogni campione rilevato. Una volta conclusa la fase di inizializzazione si restituisce un messaggio appropriato e viene cambiato il valore della variabile booleana *isSetupComplete*, per dare visibilità dell'avvenuto setup all'infuori della classe.

### 3.5.4. Metodo AnalyzeEmotions

Il metodo *AnalyzeEmotions*, riportato nella Tabella 13, analizza le informazioni contenute nell'oggetto *FaceTrackFrame* in input e le confronta con i valori ottenuti

dal setup iniziale e i *Margini di Riferimento*, per eseguire la rilevazione, restituendo come output delle stringhe che descrivono brevemente l'emozione rilevata.

```
internal String AnalyzeEmotion(FaceTrackFrame frame)
{
    if (frame.GetAnimationUnitCoefficients()[5] > this.browUpBasicValue +
    BROW_UP_GAP)
    {
        return ("Surprise!");
    }
    else if (frame.GetAnimationUnitCoefficients()[2] > this.smileBasicValue
    + SMILE_GAP)
    {
        return ("Joy!");
    }
    else
    {
        double distR = this.PointDistance(frame.GetShapePoints()[89],
        frame.GetShapePoints()[25]);
        double distL = this.PointDistance(frame.GetShapePoints()[89],
        frame.GetShapePoints()[35]);
        if (distR + BROW_LOW_GAP < browLowRBasicValue || distL + BROW_LOW_GAP
        < browLowLBasicValue)
        {
            return ("Disgust!");
        }
        else
        {
            return ("No action");
        }
    }
}
```

Tabella 13 Metodo AnalyzeEmotion

Si esegue prima un confronto sulla sorpresa e poi sulla gioia poiché sorprendersi aprendo la bocca, anche lievemente, potrebbe essere mal interpretato come una risata e quindi come gioia. Per questo motivi si controlla prima l'innalzamento delle sopracciglia, segnale inequivocabile di sorpresa.

I controlli sulle sopracciglia sono in OR e non in AND poiché ci si può disgustare anche abbassando una sola sopracciglia e non necessariamente entrambe.

Con *NoAction* viene infine indicata la mancanza di azioni particolari sul volto e la mancanza di una mimica facciale collegata alle emozioni rilevate dal programma.

## 4 Test

In questo capitolo vengono esaminate le performance dell'applicazione realizzata mettendo in risalto i risultati ottenuti con i metodi e le soglie scelte per la rilevazione delle emozioni, descritte nello specifico nel capitolo precedente.

Si sono effettuati due tipologie di test:

- Nella prima viene chiesto ad una serie di soggetti di mimare un'emozione, così da poter valutare la precisione del software sviluppato in relazione ai dati forniti su altri prodotti analoghi;
- Nella seconda gli stessi soggetti vengono sottoposti a stimoli visivi, con lo scopo di provocare un'emozione, così da poter valutare il comportamento del software in una situazione più realistica.

### 4.1. Setup sperimentale

Per lo svolgimento dei test è stata posizionata la Kinect in un ambiente con una luce naturale e diffusa, evitando quindi forti fonti luminose nel raggio d'azione del dispositivo, così da minimizzare il più possibile errori dovuti alle condizioni ambientali. Ci si è inoltre accertati che i soggetti partecipanti ai test fossero gli unici presenti nell'inquadratura per tutta la durata del singolo test, così da evitare ulteriori errori dovuti a tracciamenti di persone diverse da quella interessata.

L'applicazione ai fini dei test svolti è stata fatta girare su un computer Acer Aspire 5742G con le caratteristiche seguenti:

- Hardware:
  - Processore: Intel Core i5-480M (2.67 GHz);
  - Scheda grafica: Nvidia GeForce GT 450M;
  - Memoria installata (RAM): 4 Gb;
- Sistema Operativo:
  - Windows 10 Pro

Il dispositivo utilizzato è riportato solo per fini di completezza visto che, come già accennato precedentemente, l'applicazione è stata sviluppata per funzionare su un qualsiasi sistema operativo Windows avente installato i driver per la Kinect.



Figura 21 Esempio di setup sperimentale utilizzato nella prima sessione di test

## **4.2. Scelta dei soggetti**

Nel caso specifico dei test inclusi nella trattazione sono stati selezionati 20 soggetti, facendo attenzione che almeno il 25% dei campioni fosse composto da persone che indossano gli occhiali. Si vuole infatti studiare il comportamento dell'applicazione quando si trova a dover esaminare l'espressione di una persona quando questa porta occhiali la cui montatura che potrebbero falsare la rilevazione dei movimenti delle sopracciglia.

## **4.3. Fase di training**

Prima della fase di rilevazione vera e propria di ogni sessione di test, viene svolta una fase di training con lo scopo di far acquisire familiarità con l'applicazione ai soggetti interessati, chiedendo di eseguire le tre espressioni di interesse. Così facendo

ci si può inoltre accertare della corretta sovrapposizione tra i *Tracked Points* tracciati dall'applicazione e i corrispettivi punti sul volto del soggetto, evitando dunque errori dovuti a cause esterne all'applicazione stessa.

#### 4.4. Parametri di valutazione

AI fini dell'elaborazione i singoli risultati ottenuti sono suddivisi in tre categorie:

- *True Positive (TP)*: corretta rilevazione dell'emozione mimata;
- *False Positive (FP)*: rilevazione di un'emozione differente da quella mimata;
- *False Negative (FN)*: mancata rilevazione dell'emozione mimata.

Non si tiene conto di risultati *True Negative* (corretta rilevazione di un'espressione neutra) poiché di poco interesse ai fini della trattazione, che mira invece ad identificare emozioni particolari.

Per valutare le prestazioni dell'applicazione in entrambe le sessioni di test ci si basa su tre parametri (riportati come equazioni nella Tabella 14):

- *Accuracy*: percentuale di *True Positive* sulla totalità dei campioni rilevati (Equazione A), rappresenta la percentuale di corretta rilevazione;
- *Positive Predicted Values (PPV)*: percentuale di *True Positive* sulla totalità dei positivi rilevati (Equazione B), rappresenta la percentuale di successo sui campioni rilevati;
- *Negative Predicted Values (NPV)*: percentuale dei *False Positive* sempre sulla totalità dei positivi rilevati (Equazione C), rappresenta la percentuale di confusione dei campioni rilevati.

$Accuracy = \frac{TP}{TP + FP + FN} \times 100\%$	(A)
$PPV = \frac{TP}{TP + FP} \times 100\%$	(B)
$NPV = \frac{FP}{TP + PF} \times 100\%$	(C)

Tabella 14 Equazioni per i parametri *Accuracy*, *PPV* e *NPV*

## **4.5. Prima sessione di test**

Di seguito viene descritta nel dettaglio la prima sessione di test effettuata e vengono analizzati i risultati così rilevati.

Ai soggetti interessati viene chiesto di posizionarsi di fronte alla telecamera, al centro dell'inquadratura, e di mimare espressioni di gioia, sorpresa e disgusto, il più naturalmente possibile, evitando mimiche facciali eccessive ed innaturali.

### **4.5.1. Fase di rilevazione**

Terminata la fase di training, descritta precedentemente, viene chiesto quindi di mantenere un'espressione neutra, sempre guardando fissi l'obiettivo della Kinect, per tutta la durata del setup iniziale dell'applicazione. Una volta terminato si passa alla parte di rilevazione vera e propria. Viene chiesto ai soggetti interessati di mimare le tre espressioni che l'applicazione rileva, ognuna per tre volte e in un ordine casuale. Tra un'espressione e la successiva viene chiesto inoltre di mantenere per qualche secondo un'espressione neutra (la stessa tenuta durante la fase di setup precedente), così da separare tra loro i vari campioni ottenuti ed evitare confusione tra due espressioni successive.

Una volta mimate tutte le emozioni richieste la rilevazione viene messa in pausa e viene stampato su un file di testo, attraverso il bottone apposito, il log ottenuto, il quale verrà esaminato come descritto in seguito.

### **4.5.2. Elaborazione dei risultati**

Un file di log ottenuto dopo ogni singolo esperimento contiene al suo interno le rilevazioni effettuate e, come accennato nel capitolo relativo all'implementazione, viene indicato con "*Joy!*" la rilevazione della gioia, con "*Surprise!*" la rilevazione della Sorpresa, con "*Disgust!*" la rilevazione del disgusto e con *No Action* la rilevazione di un'emozione che differisce dalle tre precedenti, nel caso specifico dei test in esame indica un'espressione neutra.



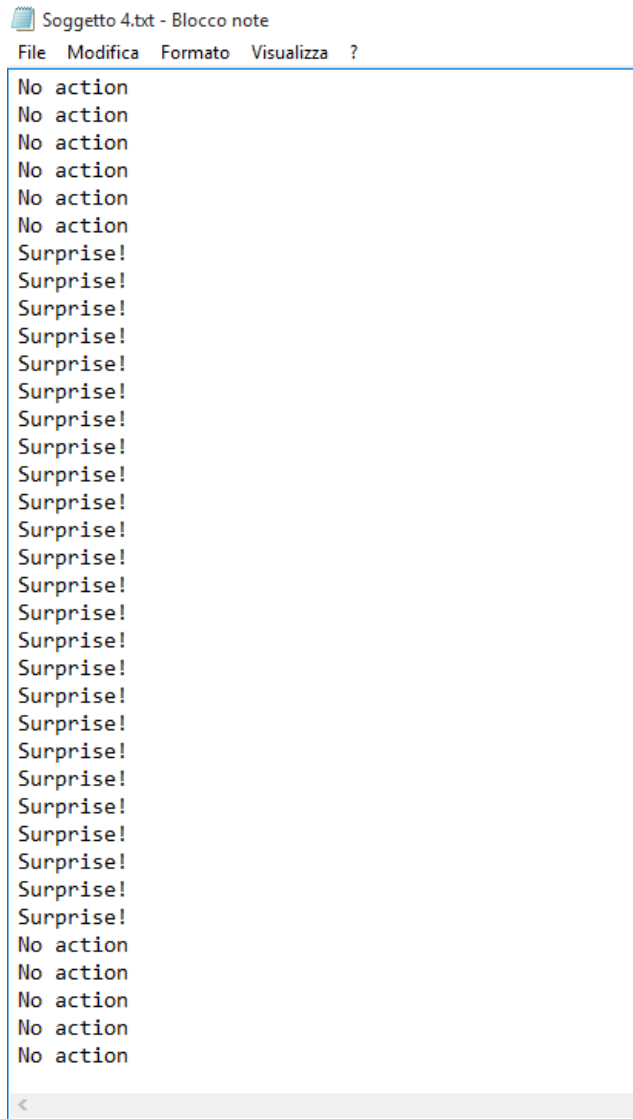


Figura 22 Esempio di un file di log

Suddetto file viene quindi esaminato per trovare le corrispondenze tra le emozioni rilevate e quelle che si è chiesto di mimare. Una lunga serie di risultati *No Action* rappresenta la fase che intercorre tra una mimica e la successiva ed è usata per separare e quindi riconoscere all'interno del file le varie emozioni rilevate.

Nel elaborare i risultati non si tiene conto di singoli risultati positivi (*Joy!*, *Surprise!*, *Disgust!*) presenti sporadicamente in una lunga serie di *No Action*, in quanto rappresentano dei falsi positivi facilmente individuabili ed eliminabili. Allo stesso modo non si tiene conto di singoli *No Action* presenti in una serie di risultati positivi. Viceversa ottenere una serie di *No Action* al posto di una serie di risultati positivi rappresenta un chiaro caso di mancata rilevazione.

#### 4.5.3. Analisi dei risultati

Come mostrato nella Tabella 15 e nella Figura 23 il software sviluppato raggiunge un'accuratezza totale del 86,7%, con un picco del 91,6% nel caso della sorpresa. Inoltre la percentuale di NPV sui campioni rilevati è inferiore al 2%.

Il software *FaceReader 6.0*, considerato il miglior software per il riconoscimento facciale attualmente in circolazione, come già accennato nel primo capitolo, vanta un'accuratezza del 81%, non fornendo però statistiche specifiche per ciascuna emozione. Non è quindi possibile effettuare un paragone assoluto tra le prestazioni di *FaceReader 6.0* e quelle del software sviluppato. I risultati ottenuti risultano comunque soddisfacenti, considerando il ristretto insieme di emozioni considerate, anche in vista di ulteriori sviluppi dell'applicazione.

EMOZIONE	TRUE POSITIVE	FALSE POSITIVE	FALSE NEGATIVE	ACCURACY	PPV	NPV	TOTALE CAMPIONI
GIOIA	51	1	8	85%	98,08%	1,92%	60
SORPRESA	55	2	3	91,6%	96,5%	3,5%	60
DISGUSTO	50	0	10	83,3%	100%	0%	60
TOTALE	156	3	21	86,7%	98,1%	1,9%	180

Tabella 15 Risultati della prima sessione di test

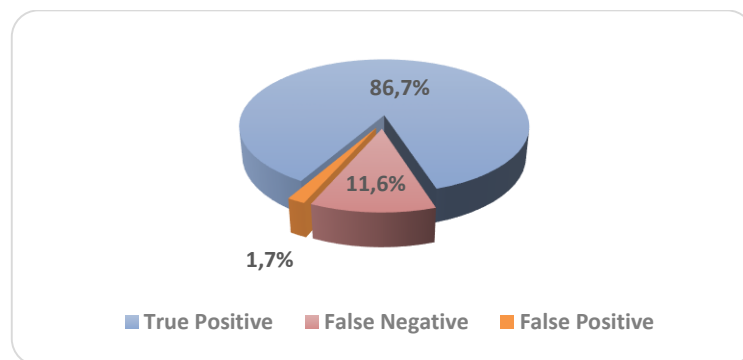


Figura 23 Grafico con le percentuali di TP, FN e FP per la prima sessione di test

#### 4.5.4. Analisi prestazioni

Come già accennato nel capitolo è di interesse studiare l'andamento dei risultati ottenuti in presenza di soggetti dotati di occhiali. In alcuni casi avviene una

sovrapposizione tra la montatura degli occhiali e le sopracciglia dei soggetti considerati, con una conseguente confusione tra le due. In questi casi risulta impossibile rilevare l'abbassamento delle sopracciglia durante la mimica dell'espressione legata al disgusto, motivo della mancata rilevazione.

Questo tipo di errore non va però ricercato in un difetto interno all'applicazione progettata, essendo ovviamente dovuto ad una condizione esterna.

Se non si tiene conto dei dati ottenuti nei 6 casi (sui 20 totali) in cui le persone esaminate risultano con questa condizione, i valori contenuti nella Tabella 15 si modificano come indicato nella Tabella 16.

EMOZIONE	TRUE POSITIVE	FALSE POSITIVE	FALSE NEGATIVE	ACCURACY	PPV	NPV	TOTALE CAMPIONI
GIOIA	38	0	4	90,4%	100%	0%	42
SORPRESA	38	2	2	95%	95%	5%	42
DISGUSTO	39	0	3	92,8%	100%	0%	42
TOTALE	115	2	9	91,27%	98,3%	1,7%	126

Tabella 16 Risultati della prima sessione di test senza tener conto di soggetti con occhiali

I risultati ottenuti dimostrano un notevole aumento generale dell'accuratezza del programma, soprattutto nella specifica rilevazione del disgusto, a conferma dell'ipotesi secondo la quale la presenza di occhiali su un soggetto influisce negativamente sulla rilevazione generale, in particolare dell'espressione legata al disgusto.

#### 4.6. Seconda sessione di test

Di seguito viene descritta nel dettaglio la seconda sessione di test effettuata e viene fornita una valutazione dei risultati ottenuti, anche in relazione a quelli della sessione precedente.

#### **4.6.1. Scelta degli stimoli**

Non è possibile trovare stimoli universali per provocare le emozioni riconosciute, essi sono infatti fortemente dipendenti dalla provenienza culturale e possono variare drasticamente da soggetto a soggetto, anche all'interno dello stesso contesto sociale. Si sono scelti quindi stimoli ritenuti adeguati per l'insieme di soggetti presi in esame. Per quanto riguarda la gioia si sono scelte immagini con didascalie divertenti, con lo scopo di provocare sorrisi o risate.

Studi precedenti [11] dimostrano che il disgusto può essere facilmente causato dalla vista di cibo estraneo alla propria cultura di appartenenza, si sono dunque scelti stimoli di questa natura.

Come già accennato nel Capitolo 1, la sorpresa è difficilmente rilevabile vista la sua breve durata e si manifesta sempre in concomitanza di altre emozioni. Si sono quindi scelti stimoli con lo scopo di provocare una reazione di incredulità nei soggetti, poiché caratterizzata dalla stessa mimica facciale della sorpresa, in modo da evitare rilevazioni di altre emozioni insieme alla sorpresa stessa.

Sono stati anche inseriti una serie di stimoli neutri, con lo scopo di non causare nessuna espressione particolare nei soggetti.

Con lo scopo di rilassare il soggetto ed evitare sovrapposizione tra le varie emozioni, tra uno stimolo ed il successivo è stata riprodotta per 5 secondi un'immagine composta da una tonalità di grigio uniforme (R=128, G=128, B=128),

#### **4.6.2. Fase di rilevazione**

Ai soggetti interessati viene chiesto di posizionarsi davanti ad uno schermo sul quale verrà riprodotto un video contenente gli stimoli selezionati. Viene sottolineata la necessità di guardare lo schermo e non la Kinect per tutta la durata del test, soprattutto durante il setup iniziale dell'applicazione. Ci si assicura inoltre, come nella sessione precedente, che i soggetti interessati siano al centro dell'inquadratura della Kinect.

Una volta concluso il setup viene avviato il video con gli stimoli. Terminato quest'ultimo l'applicazione viene messa in pausa attraverso il bottone apposito ed i

risultati vengono salvati su un file di testo come già avvenuto nella sessione di test precedente.

#### **4.6.3. Elaborazione dei risultati**

Il file di log ottenuto in questo esperimento presenta le stesse caratteristiche di quello descritto nel Paragrafo 4.4.3.

L'unica differenza è rappresentata da una lunga serie di *No Action* più lunga delle altre, che ci si aspetta di rilevare in corrispondenza degli stimoli neutri.

#### **4.6.4. Analisi dei risultati**

I risultati ottenuti si dimostrano in linea con le aspettative, anche in relazione alle considerazioni già fatte durante la trattazione.

Come mostrato nella Tabella 17 il calo delle prestazioni generali è sicuramente da ricercarsi nella difficoltà riscontrata nel provocare le emozioni, piuttosto che chiedere di mimarle. Molti soggetti non si sono infatti affatto sentiti disgustati, divertiti o sorpresi dagli stimoli proposti.

Di notevole interesse sono i risultati ottenuti per quanto riguarda la sorpresa. Il notevole calo delle prestazioni ottenuto in questo caso specifico conferma infatti la difficoltà nell'elicitare e rilevare tale emozione.

In questa sessione di test, a differenza della precedente, sono stati inoltre raccolti campioni riguardanti la rilevazione dell'emozione neutra in risposta al relativo stimolo. I risultati ottenuti dimostrano una buona *Accuracy* e quindi una bassa tendenza dell'applicazione a rilevare un'emozione quando questa non è effettivamente presente sul volto.

EMOZIONE	TRUE POSITIVE	FALSE POSITIVE	FALSE NEGATIVE	ACCURACY	PPV	NPV	TOTALE CAMPIONI
GIOIA	47	0	13	78,33%	100%	0%	60
SORPRESA	24	12	24	40%	66,67%	33,33%	60
DISGUSTO	43	3	14	71,66%	93,47%	6,53%	60
NEUTRO	56	4	ND	93,3%	93,3%	6,7%	60
TOTALE	170	19	41	70,83%	89,94%	10,6%	240

Tabella 17 Risultati della seconda sessione di test

Come discusso nei paragrafi precedenti i *False Negative* sono definiti come occorrenze di *No Action* nel file di log al posto dell'emozione desiderata. Nel caso della rilevazione dell'espressione neutra il risultato cercato è proprio *No Action*, il che ha portato ad, in questo caso specifico, un'impossibilità di fornire una stima precisa dei valori *False Negative*, indicata attraverso la sigla ND (Non Disponibile) nella Tabella 17.

## Conclusione e Sviluppi futuri

In questo lavoro di tesi si sono esaminati i passaggi fondamentali per lo sviluppo di un'applicazione in grado di determinare l'emozione di un utente, con lo scopo di valutarne la *Quality of Experience*.

Particolare attenzione è stata data agli studi effettuati in merito al riconoscimento delle emozioni, proprio con lo scopo di identificare una buona tecnica per il rilevamento dell'emozione dell'utente.

L'applicazione sviluppata presenta una buona percentuale di successo nel caso in cui venga richiesto all'utente di mimare una certa espressione, attestandosi sullo stesso livello di altri software presenti nello stato dell'arte.

Nel caso di test in condizioni reali, durante il quale i soggetti sono stati sottoposti a stimoli con lo scopo di provocare un'emozione, l'applicazione presenta cali di rendimento. Questi cali di precisione sono però attesi e giustificati dagli studi effettuati, assimilabili proprio con la difficoltà di causare un'emozione specifica in un soggetto: si sono riscontrati cali maggiori proprio nella rilevazione delle emozioni più difficili da elicitare.

In quest'ultimo caso non è inoltre possibile fare paragoni accurati con altri software, in quanto non è presente un'esaustiva documentazione da parte degli sviluppatori per produrre una statistica attendibile.

Si prevedono nel futuro nuove campagne di test con un maggior numero di soggetti, così da poter raffinare il sistema sviluppato anche in vista di un utilizzo in contesti più pratici e di più alto valore aggiunto. È infatti lecito pensare all'utilizzo di questa applicazione in casi in cui è necessario rilevare la *Quality of Experience*, in maniera non invasiva, di soggetti che guardano un film, che partecipano ad un colloquio o che giocano ad un videogioco. La possibilità di servirsi dell'applicazione è indubbiamente vasta ed è limitata esclusivamente dall'immaginazione. Si punta molto sull'aspetto videoludico, in quanto l'hardware scelto per lo sviluppo, il sensore Kinect, è un accessorio di una delle console di maggior diffusione al momento della trattazione. Va però ricordato che l'utilizzo del sensore non è vincolato alla console, in quanto ne esistono versioni progettate per interfacciarsi con un normale computer.

È infine necessario sottolineare che ulteriori sviluppi dell'applicazione sono strettamente legati proprio alla tecnologia usata. Molte delle limitazioni imposte durante la trattazione sono infatti dovute a mancanze legate al sensore, causate soprattutto dalla “giovane età” di quest’ultimo e al fatto che lo si sta usando in ambito di ricerca da solo qualche anno. Un miglioramento della Kinet non potrà fare altro che portare ad un perfezionamento di ugual portata nell'applicazione sviluppata.



## Bibliografia

- [1] In che modo Google utilizza il riconoscimento di pattern per comprendere le immagini  
Available:  
<https://www.google.it/intl/it/policies/technologies/pattern-recognition/>
- [2] Il riconoscimento facciale di Windows 10 supera la prova dei gemelli  
Available:  
<http://www.focus.it/tecnologia/digital-life/il-riconoscimento-facciale-di-windows-10-supera-la-prova-dei-gemelli>
- [3] DeepFace, l'algoritmo di Facebook che riconosce i volti degli utenti  
Available:  
<http://urbanpost.it/deepface-lalgoritmo-di-facebook-che-riconosce-i-volti-degli-utenti>
- [4] È made in China il primo bancomat con il riconoscimento facciale  
Available:  
<http://www.cinaforum.net/e-made-china-il-porimo-bancomat-riconoscimento-facciale/>
- [5] When Algorithms Grow Accustomed to Your Face  
Available:  
[http://www.nytimes.com/2013/12/01/technology/when-algorithms-grow-accustomed-to-your-face.html?\\_r=0](http://www.nytimes.com/2013/12/01/technology/when-algorithms-grow-accustomed-to-your-face.html?_r=0)
- [6] Teaching Cars To Feel Your Pain  
Available:  
<http://spectrum.ieee.org/tech-talk/transportation/human-factors/teaching-cars-to-feel-your-pain>

- [7] Tecnologie per la lettura automatica delle emozioni: quali conseguenze?  
Available:  
[http://www.agendadigitale.eu/egov/1041\\_tecnologie-per-la-lettura-automatica-delle-emozioni-quali-conseguenze.htm](http://www.agendadigitale.eu/egov/1041_tecnologie-per-la-lettura-automatica-delle-emozioni-quali-conseguenze.htm)
- [8] S. J. Haggbloom, The 100 Most Eminent Psychologists of the 20th Century. Review of General Psychology, 2002.
- [9] P. Ekman, I Volti della Menzogna. Gi indizi dell'inganno nei rapporti interpersonali, 2014.
- [10] 8 Biggest Myths About Lying According to the Best Human Lie Detector in the World  
Available:  
[http://www.huffingtonpost.com/amy-morin/the-8-biggest-myths-about-lying-according-to-the-best-human-lie-detector\\_b\\_7568892.html](http://www.huffingtonpost.com/amy-morin/the-8-biggest-myths-about-lying-according-to-the-best-human-lie-detector_b_7568892.html)
- [11] P. Ekman, Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life, 2003.
- [12] P. Ekman and W. Friesen, Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, 1978.
- [13] Cohn-Kanade (CK and CK+) database Download Site  
Available:  
<http://www.consortium.ri.cmu.edu/ckagree/>
- [14] T. Kanade, J. F. Cohn and Y. Tian, Comprehensive database for facial expression analysis, 2000.
- [15] P. Ekman, W. V. Friesen, Constants across cultures in the face and emotion, 1971.
- [16] Available:  
<http://neurologicalcorrelates.com/wordpress/2008/09/01/book-review-human-the-science-behind-what-makes-us-unique-by-michael-s-gazzaniga-and-some-on-line-deception-detection-tests/>

- [17] Available:  
<http://www.theguardian.com/lifeandstyle/2009/may/12/psychology-lying-microexpressions-paul-ekman>.
- [18] Available:  
<http://www.theatlantic.com/health/archive/2015/02/hard-feelings-sciences-struggle-to-define-emotions/385711/>
- [19] Microsoft Kinect  
Available:  
[https://it.wikipedia.org/wiki/Microsoft\\_Kinect](https://it.wikipedia.org/wiki/Microsoft_Kinect).
- [20] Kinect for Windows Sensor Components and Specifications  
Available:  
<https://msdn.microsoft.com/en-us/library/jj131033.aspx>.
- [21] M. Caggiano, Gesture Recognition per applicazioni domotiche mediante sensore di visione 3D, Tesi di Laurea, 2012.
- [22] Method and System for Object Reconstruction  
Available:  
<https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2007043036>.
- [23] Kinect for Xbox One  
Available:  
[https://en.wikipedia.org/wiki/Kinect\\_for\\_Xbox\\_One](https://en.wikipedia.org/wiki/Kinect_for_Xbox_One)
- [24] All About Kinect  
Available:  
<http://www.i-programmer.info/babbages-bag/2003-kinect-the-technology-.html>
- [25] Skeletal Tracking  
Available:  
<https://msdn.microsoft.com/en-us/library/hh973074.aspx>
- [26] What's New  
Available:  
[https://msdn.microsoft.com/en-us/library/jj663803.aspx#SDK\\_1pt8](https://msdn.microsoft.com/en-us/library/jj663803.aspx#SDK_1pt8)

- [27] Microsoft.Kinect.Toolkit.FaceTracking Component  
Available:  
<https://msdn.microsoft.com/en-us/library/jj131022.aspx>
- [28] Face Tracking  
Available:  
<https://msdn.microsoft.com/en-us/library/jj130970.aspx>
- [29] J. Hall, P. Perry, S. Kean, Meet the Kinect. An Introduction to Programming Natural User Interfaces, 2012.
- [30] MacRobert Award  
Available:  
[https://en.wikipedia.org/wiki/MacRobert\\_Award](https://en.wikipedia.org/wiki/MacRobert_Award)
- [31] Microsoft Xbox 360 Kinect al top delle vendite a novembre  
Available:  
[http://www.affaritaliani.it/mediatech/microsoft\\_xbox\\_360\\_kinect101210.html](http://www.affaritaliani.it/mediatech/microsoft_xbox_360_kinect101210.html)
- [32] Introduction to WPF  
Available:  
<https://msdn.microsoft.com/en-us/library/aa970268%28v=vs.100%29.aspx>
- [33] XAML Overview  
Available:  
<https://msdn.microsoft.com/en-us/library/cc189036%28v=vs.95%29.aspx>
- [34] Working with App.xaml  
Available:  
<http://www.wpf-tutorial.com/wpf-application/working-with-app-xaml/>
- [35] Mysteries of Kinect for Windows Face Tracking output explained  
Available:  
<http://blogs.msdn.com/b/kinectforwindows/archive/2014/01/31/clearing-the-confusion-around-kinect-for-windows-face-tracking-output.aspx>.
- [36] CANDIDE – a parametrized face  
Available:  
<http://www.icg.isy.liu.se/candide/main.html>