

# 接客エージェントへ生物感を付与する視線制御モデル

## A gaze control model making an agent look like alive

永吉秋平<sup>1\*</sup> 奥岡耕平<sup>1</sup> 北川遼<sup>1</sup> 岨野太一<sup>1</sup> 今井倫太<sup>1</sup>

Shuhei Nagayoshi<sup>1</sup> Kohei Okuoka<sup>1</sup> Ryo Kitagawa<sup>1</sup> Taichi Sono<sup>1</sup> Michita Imai<sup>1</sup>

<sup>1</sup> 慶應義塾大学

<sup>1</sup> Keio University

### Abstract:

本研究はユーザへ接客を行うエージェントに焦点を当て、エージェントがユーザの周囲の環境へ反応する機構を実装することで、エージェントに対する生物的な感覚をユーザに与え、サービスがもたらすユーザエクスペリエンスの向上を試みる。エージェントが行うユーザの周囲の環境への反応として、エージェントの視界に現れた人物に対し優先度を考慮した視線の制御を実装した。

## 1 序論

情報科学の発展により CG エージェントやロボットといったエージェントが接客を行う機会が増えている。例えば SoftBank のペッパー<sup>1</sup> が飲食店などで受付を担うなど実際にエージェントが活躍する例がある。エージェントとユーザのインタラクションにおいてユーザがエージェントに対して抱く印象はユーザエクスペリエンスを向上させる上で重要である。エージェントがユーザの接客を行う際には、単に接客を行うだけではなく、エージェントに親しみやすい印象をユーザに与えることが、ユーザエクスペリエンスの向上にとって大切なファクターとなる。

エージェントに対して親しみやすい印象を与えるために、従来研究では様々なアプローチが提案されている。目良 [1] は、エージェントの感情を付与することで親しみやすいコミュニケーションの実現を行っている。福田ら [2] は嗜好や感情を提案し、親しみやすさを向上させることを示している。

人間は無生物のオブジェクトよりも生物感を感じる相手に対して寛容であるという研究 [3] がある。エージェントに生物感を持たせるための要因として、随伴性と呼ばれる、他者の振る舞いに対して即座に反応する性質に着目した研究がある。[4] では、ロボットの随伴性とインタラクションの複雑さに着目してロボットの生物らしさについての研究を行っている。

随伴性行動によって生物感を付与する既存研究において、エージェントに随伴性行動を実装することによ

て生物感を付与することの重要性は示されている。しかし、既存研究の多くは特定のタスクを想定しない汎用的なエージェントを対象としており、特定のタスクを想定した実装について検討していない。

そこで本論文では、接客エージェントに着目し、接客タスクを考慮した随伴性行動を実現するための行動モデルを提案する。

## 2 背景

### 2.1 関連研究

#### 2.1.1 接客エージェント

人間とコミュニケーションをとるエージェントの普及が進み、多種多様なシチュエーションでのエージェントの活用が見込まれている。中でもコミュニケーションロボットや接客を行うエージェントなど、親しみやすいエージェントの需要が増加している [1, 5]。

石井らの研究 [6] では適応的会話制御を実装したエージェントの評価実験として研修中の新人販売員エージェントの接客を想定した実験を行っている。また、出見世らの研究 [7] ではオンラインショッピングに焦点を当て、自律的かつ適応的な接客を行うエージェントの提案を行っている。久保ら [8] は接客ロボットが決められたシナリオにしか対応できないという問題を指摘し、遠隔操作と自律運動を組み合わせることで接客の質を高めようとしている。

\*連絡先：(慶應義塾大学)  
(神奈川県横浜市港北区日吉3丁目14-1)  
E-mail:nagayoshi@ailab.ics.keio.ac.jp

<sup>1</sup>Pepper (ペッパー) — ロボット — ソフトバンク  
(<https://www.softbank.jp/robot/pepper/>)

### 2.1.2 インタラクションの質の向上

エージェントとユーザの関わりについて調査を行い、エージェントとのインタラクションの質を向上させようと試みている研究は数多くある。[3]では無生物のオブジェクトよりも生物的な相手に対して人間は寛容になれるとしている。

エージェントに対する親しみをユーザに持たせる既存研究は数多く見られる。目良 [1] は、エージェントの感情を付与することで親しみやすいコミュニケーションの実現を目指している。竹内ら [5] はエージェントに感情を持たせる感情生成モデルを提案している。伊藤ら [9] はロボットの内部状態を代表して感情生成モデルを提案し、気分の変化をロボットの個性を評価した。福田ら [2] は嗜好や感情を更新するエージェントを提案した。

### 2.1.3 生物感の付与と随伴性

随伴性とは他者の振る舞いに対して即座に反応する性質である。エージェントに生物感を持たせるために随伴性に着目する研究が見られる。[4]では、ロボットの随伴性とインタラクションの複雑さに着目してロボットの生物らしさについての研究を行っている。滝本らの研究 [10] では優先度と重み付き平均化の2段階で振る舞いの統合を行う随伴性アーキテクチャを提案している。

### 2.1.4 視線の役割

随伴性を踏まえて生物感をエージェントに付与する上で、重要な役割を担っている1つが視線である。Kendon [11] によると視線の機能には、対象の視覚情報を取得する「監視機能」、会話における発話権の調整をする「調整機能」、その他の情報を相手に伝える「表現機能」がある。表現機能の具体例としては、視線によって表現者の意図や感情、対象への関心度を伝えるといったものがある。深山らの研究 [12] では印象というタームを人格の予測や好悪・善悪などの評価から成る、対象人物の心的イメージと定義している。そして印象を凝視量、凝視持続時間、非凝視時視線位置を視線をコントロールするパラメータとして、視線がエージェントの印象に与える影響を調査している。これは Kendon [11] の視線の表現機能に着目したものである。

## 2.2 本研究での取り組み

エージェントとユーザのインタラクションにおいてユーザがエージェントに対して抱く印象はユーザエクスペリエンスを向上させる上で重要である。エージェントがユーザの接客を行う際には、単に接客を行うだ

けではなく、エージェントに親しみやすい印象をユーザに与えることが、ユーザエクスペリエンスの向上にとって大切なファクターとなる。しかしエージェントの印象を親しみやすいものにするを目的として生物感をエージェントに付与する研究は、接客エージェントで検討されていない。

そこで本研究では、接客エージェントに焦点を当て、エージェントに生物感を付与するために周囲の環境に反応する機能を実装した。周囲の環境への反応として、視線の表現機能に着目してエージェントの視界に映る人物に視線を向ける振る舞いを採用した。これは深山らの研究 [12] によりエージェントの視線がエージェントの印象に影響を与えることが判明しているためである。本論文で採用した視線の振る舞いは Kendon [11] の視線の3つの機能のうち視線の表現機能にあたる。

## 3 提案

接客エージェントが視線を向けるユーザを決めるために、本提案モデルでは接客エージェントが認識している各ユーザに優先度を導入する。今回優先度を算出する式を提案するに当たって考慮した条件が

- 人物が新しくエージェントの視界に現れたこと
- 人物が継続してエージェントの視界に映り続けていること
- 接客を行っている対象のユーザであること

である。これは、エージェントがユーザの周囲の環境に反応すること及び接客エージェントという特性を考慮したためである。

$$P_X = P_{X_{new}}(t_x) + P_{X_{gaze}}(t_x) + P_{center} \quad (1)$$

式1で人物  $X$  に対して線を向ける優先度  $P_X$  を定義する。 $t_x$  は人物  $X$  が画面に現れた時間を  $t_x = 0$  とする時間を表す変数である。視線を向ける人物は  $P_X$  を最大にする  $X$  とする。

$$P_{X_{new}}(t_x) = \frac{1}{1 + e^{\alpha t_x + \beta}} \quad (2)$$

$$P_{X_{gaze}}(t_x) = \frac{\delta}{\gamma t_x + 1} \quad (3)$$

$$P_{center} = \begin{cases} C & (X \text{ が最も中心に近い}) \\ 0 & (otherwise) \end{cases} \quad (4)$$

式3,4で  $P_{X_{new}}$ ,  $P_{X_{gaze}}$ ,  $P_{center}$ , を定義する。 $\alpha$  は新規係数,  $\beta$  は調整定数,  $\gamma$  は凝視係数,  $\delta$  は凝視負荷係

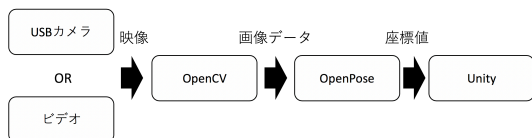


図 1: システム概略図

数,  $C$  は対応優先度と名付け, 実装に応じて変化させるパラメータであり今回の環境では  $\alpha = 0.09, \beta = 10, \gamma = 0.05, \delta = -2, C = 0.2$  として調整を行った.

$P_{X_{new}}$  は新しく顔を覗かせたことによる優先度で時間を変数とする下に有界な減少関数である. 最初に顔を覗かせてから時間が経過していくと段々優先度が下がっていくことを表している.

$P_{X_{gaze}}$  は同じ人物が顔を覗かせ続けていることによる優先度で時間を変数とする上に有界な増加関数である. ずっとエージェントの視界の中に入り続けていると次第に優先度が上がっていくことを示している.

$P_{center}$  は人物  $X$  が最も中心に近い座標に存在する人物である場合のみ加算される定数である. 接客というタスクにおいて対象となる人物が現在対応している人物であるかを考慮する必要がある. 対応している人物に視線を向けるために対応している人物の優先度に定数を加算する. ここで, 対応すべき人物か否かを判断する方法として, エージェントの視界の中央に最も近い人物を対応すべき人物として優先度の加算を行う方法を選択した.

3つの優先度の和を取ることで人物  $X$  に対する優先度とする. 新しく顔を覗かせることに寄って  $P_{X_{new}}$  の値が瞬間的に大きくなり新しく現れた人物の方を向くことを実現する. また,  $P_{center}$  の値によって接客を行うべき人物に対して他の人物よりも優先して視線を向けさせることが出来る.

## 4 実装

全体のシステムの概略図を図 1 に示す.

映像の取得は Python のライブラリ OpenCV[13] を用いた. OpenCV で取得した映像は画像データとして同じく Python のライブラリである

OpenPose[14] に渡され, OpenPose で人物の検出や座標の取得といった処理が行われる. OpenPose は単眼カメラ画像を用いて人物のスケルトンを検出することができるライブラリである.

接客エージェントにおいて着目すべき人物の部位は顔であるということ, 複数の座標値の平均値を算出すると対象となる関節が画面内に写っているか否かによって座標値に大きく違いが出てしまう. そのため今回の実装では 1 人の人物から得られる複数の関節の座標

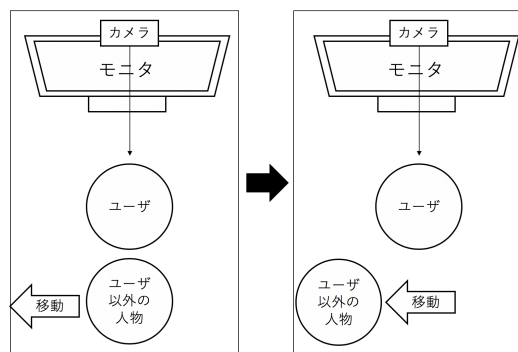


図 2: モニタとユーザの配置

のうち, 鼻にあたる関節の座標を人物の位置する座標として用いた.

エージェントの表示と視線の移動はゲームエンジン Unity を用いた.

## 5 動作例

実際にエージェントが USB カメラを通した映像から人物の座標に応じて顔を動かしている様子を示す.

動作時のモニタと対応するユーザ及び第三者の配置を図 2 に示す. モニタの前にエージェントが接客を行うユーザが立ち, その後ろに別の人物が隠れるように立つ. 後ろに立った人物がユーザの横から顔を覗かせることでエージェントが顔を覗かせた人物に反応することを確認した.

エージェントが写っている人物に追従して左右に視線を動かす様子を 3 に示す. 図では人物と反対の方向に顔を動かしているように見えるが, これは USB カメラの映像は実際の動きとは左右反対の映像を映し出すためである. 映り込んでいる人物が 1 人だけである場合は当該人物の優先度が常に一番高くなるため追従して視線を動かす. これは, 接客エージェントが, エージェントの視界内に 1 人だけの場合に対応すべき人物のほうを常に注視し続けるということに相当する.

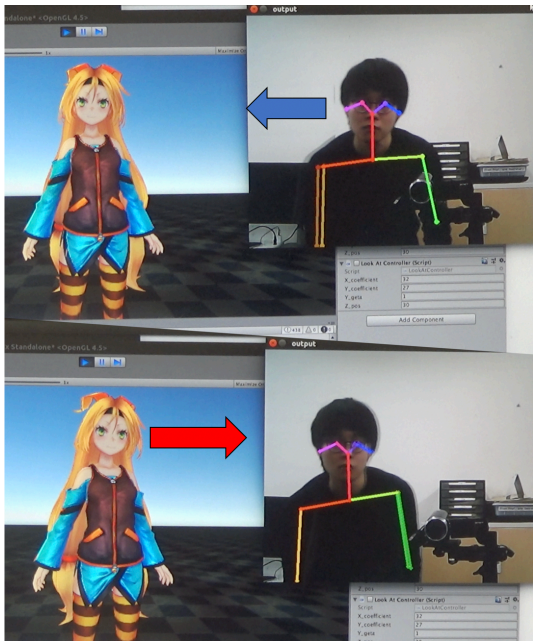


図 3: 右に移動する人物に追従するエージェント

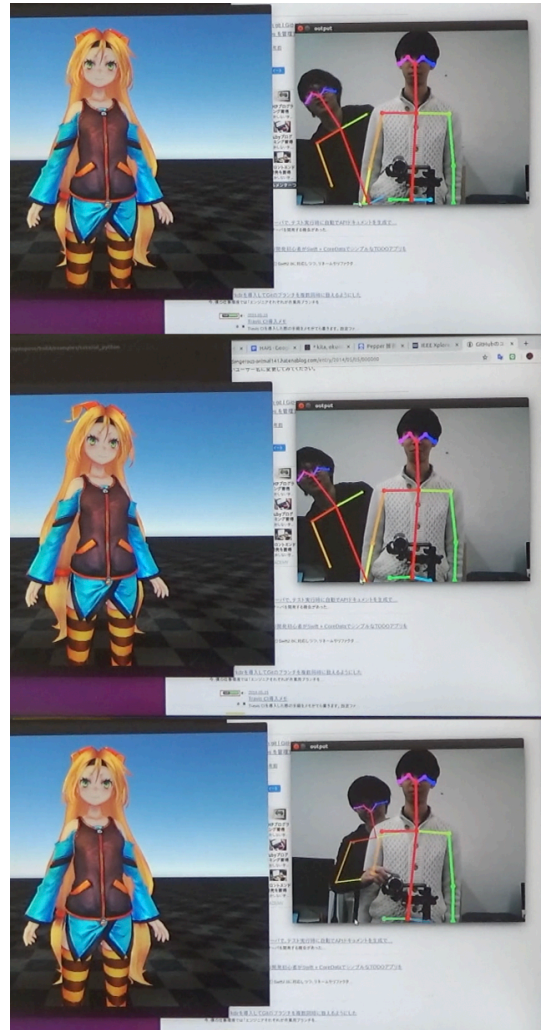


図 5: 応対者へ視線を戻すエージェント

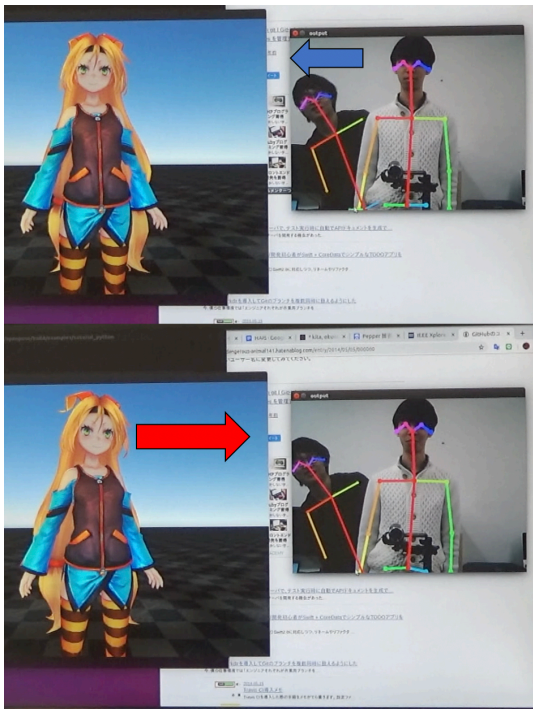


図 4: 新しい人物に追従するエージェント

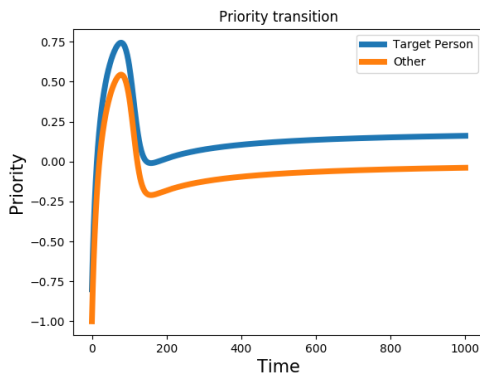


図 6: 同時に 2 人が現れた場合の優先度の遷移

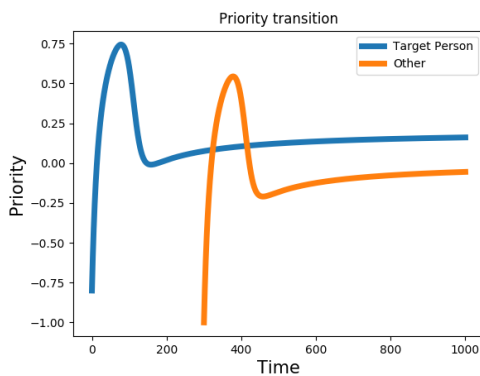


図 7: 後から 1 人が現れた場合の優先度の遷移

接客を行っている相手以外の人物が視界に現れたときにエージェントがそちらに顔を向ける様子を図 4 に、接客を行っている相手以外の人物が現れてからしばらくして再び応対すべき人物に視線を戻す様子を図 5 に示す。新しい人物が現れると元々映っていた人物の優先度よりも新しく現れた人物の優先度が高くなりエージェントはそちらに視線を向ける。しかし時間が経つと優先度が下がるため、接客する人物に視線を戻す。今回の実装において接客すべき人物は画面の中央に近い位置に座標が検出されている人物とした。

優先度の推移をグラフにしたものを図 6,7 に示す。図 6,7 の優先度は接客相手とそうでない人物の優先度を示す。図 6 は 2 人の人物が同時にエージェントの視界に現れたとしている。図 7 は 2 人目の人物が遅れてエージェントの視界に現れたとしている。

## 6 今後の課題

### 6.1 接客タスクの実装

本論文では接客エージェントに焦点を当て、ユーザとの接客を行いつつユーザの周囲の環境に反応することで接客エージェントに生物感を付与することを提案した。今回の実装では視線を動かすことでユーザの周囲に反応するモジュールを実装したが、実際に接客タスクを同時に行った場合の影響を考慮する必要がある。例えば、タスクとは無関係な動作が多くなると本来のタスクである接客を妨げる可能性がある。接客エージェントに対する印象の向上を目的としたタスク外の環境への反応を本来の目的であるタスクを妨げないように調整する必要がある。

### 6.2 多様なシチュエーションへの対応

#### 6.2.1 人物の出現パターン

本研究で提案した式 1 は新しく現れたことによる優先度と映り続けていることによる優先度と応対すべき人物かどうかの優先度の和をとっている。図 6 が示すように優先度の和は最終的には上に有界になる。最終的な優先度が上に有界であるため、十分時間が経つと応対相手の優先度は他の人物の優先度より常に高くなる。つまり応対すべきでない人物が長時間映り続けていても、応対すべき人物の優先度が常に高い。そのため、応対相手以外の人物が常にエージェントの視界に移り続けている場合でも、その人物に視線を向けることができない。

ATM の画面でユーザに操作の案内をする接客エージェントの場合、ユーザの後ろに並んで人間が横から顔を覗かせた場合、エージェントが一度視線を向けるだけでは不十分である。後ろに並んで人間がずっと覗き続けていた場合に再び覗いている人間に視線を向ける動作が自然であると考えられる。今回の実装ではそのようなシチュエーションに対応できない。

#### 6.2.2 動作のパターン

本研究におけるエージェントの反応は人物の顔の認識によって行われる。しかし実際に接客を行う場面では顔以外に人物の振る舞いも認識して何らかの反応をすることが望ましい。前項で述べた ATM での操作案内を行う接客エージェントの場合、操作しているユーザの後ろに並んでいる人物が不審な動きを見せたときに反応するといったシチュエーションが考えられるが、顔のみを考慮した現在のシステムでは不十分である。身体

動きを認識する機能及び、認識した身体の動きを考慮したモデルの設計が必要である。

### 6.3 複数の人物間での優先度の調停

本研究では複数の人物間での優先度に基づいた視線の移動を実装した。式4で示しているように優先度は中央に近い人物をを対応すべきユーザとし、それ以外のユーザ全てを対応していないユーザとしている。

しかし複数人をエージェントが認識している場合、対応すべきユーザ以外の間でも優先度に区別をつけるべき場合が考えられる。例えば接客しているユーザの後ろに複数人並んでいる場合、列の順序に基づいて優先度を割りあてるといったことが考えられる。

## 7 結論

本論文では、ユーザの周囲の環境へ反応する機構を実装することで、接客エージェントの振る舞いに生物感を与え、サービスがもたらすユーザエクスペリエンスの向上を試みた。

接客エージェントが行うユーザの周囲の環境への反応として優先度を用いた視線の制御モデルを提案した。優先度を式に従って算出し、最も高い優先度を持つ人物の方向へ接客エージェントが視線を向ける。

今後の課題として、接客タスクの実装と接客タスクと周囲への反応の兼ね合い、様々なシチュエーションへの対応、複数の人物間での優先度の調停が必要であると考えられる。

## 参考文献

- [1] 目良和也, 市村匠, 黒澤義明, 竹澤寿幸. 情緒計算手法と心的状態遷移ネットワークを用いた音声対話エージェントの気分変化手法. 知能と情報, Vol. 22, No. 1, pp. 10–24, 2010.
- [2] 福田和彰, 加藤昇平. 親しみやすい感性対話エージェントの実現に向けた動的嗜好更新手法. 電気学会論文誌C (電子・情報・システム部門誌), Vol. 138, No. 12, pp. 1485–1492, 2018.
- [3] Clark Elliott. 人間とコンピュータの間の双方向感情的コミュニケーションの構成要素-感情と人格に関する広範な基本モデルを用いて. 認知科学, Vol. 1, No. 2, pp. 16–30, 1994.
- [4] Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. How contingent should a lifelike robot be? the relationship between contingency and complexity. *Connection Science*, Vol. 19, No. 2, pp. 143–162, 2007.
- [5] 竹内将吾, 酒井あゆみ, 加藤昇平, 伊藤英則. 対話者好感度に基づく感性会話ロボットの感情生成モデル. 日本ロボット学会誌, Vol. 25, No. 7, pp. 1125–1133, 2007.
- [6] 石井亮, 中野有紀子ほか. ユーザの注視行動に基づく会話参加態度の推定—会話エージェントにおける適応的会話制御に向けて. 情報処理学会論文誌, Vol. 49, No. 12, pp. 3835–3846, 2008.
- [7] 出見世有紀, 須田飛志, 中沢実, 服部進実. 自律的接客行為を学習するインタラクティブショッピングシステム. 電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎, Vol. 97, No. 331, pp. 23–28, 1997.
- [8] 久保隆宏, 中山光樹ほか. 自動応答と遠隔操作を組み合わせた接客業務の効率化. *SIG-SLUD*, Vol. 5, No. 02, pp. 47–48, 2016.
- [9] 千加伊藤, 昇平加藤, 英則伊藤. 感性会話ロボットの性格付けとその感性評価. 日本感性工学会論文誌, Vol. 8, No. 3, pp. 899–906, 2009.
- [10] Yusuke Takimoto, Komei Hasegawa, Taichi Sono, and Michita Imai. A simple bi-layered architecture to enhance the liveness of a robot. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*, pp. 2786–2792. IEEE, 2017.
- [11] Adam Kendon. Some functions of gaze-direction in social interaction. *Acta psychologica*, Vol. 26, pp. 22–63, 1967.
- [12] 深山篤, 大野健彦, 武川直樹, 澤木美奈子, 萩田紀博. 擬人化エージェントの視線による印象操作. インタラクシオン, Vol. 2002, pp. 93–100, 2002.
- [13] Gary Bradski and Adrian Kaehler. *Opencv*. Dr. Dobb's journal of software tools, Vol. 3, , 2000.
- [14] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1302–1310. IEEE, 2017.