

アテンション・ブリーズ

- 注意の計測，効果とその応用 -

主幹研究員 / メディア情報研究グループリーダー

大和 淳司

人は日常の活動の中で、膨大な量の情報を受け取り、それに基づいて行動を重ねている。人間が情報を受け取る入り口は、視覚・聴覚・嗅覚・味覚・触覚(一般に五感と呼ばれる)が主な物だ。その中でも視覚と聴覚は、離れていても感じられるため、人をとりまく全方向からの情報が押し寄せ、これらの全てを受け取り即時に処理することは、人間にとって大きな負荷になる。そのため、人間は(動物もだが)外界のすべてから万遍なく情報を受け取るのではなく、重要なところ、興味のあるところから集中的に受取る仕組みを持っている。これが「注意(attention)」を向けるということであり、注意を向けた先から得たごくわずかな情報だけを処理することで、人間は、日々の行動を効率よく進められる。

注意のなかで、特に視覚的な注意は、情報の受信と同時に発信源ともなっていることが他には見られない特徴である。視線を向けるという行為は、自らが視線を向けた先から情報を得ると同時に、そこに注意を向けているという事実を他人にも知らしめている。人を見る場合、見られた相手は自分が見られていることに極めて敏感であるし、人の視線が向いている先がどこかを検知する能力にもたけている。

我々は人間のこのような特性をうまく利用して、人と機械とのインタラクション、機械を介した人と人とのコミュニケーションをより自然に気持ちよく、また効率のよいものにしていこうと考えている。

注意の効果

二人の人が注意を互いに向け合えば、アイコンタクトが成立する。このとき、人は自分が見られていることを意識すると同時に、見ている自分をも強く意識する。これにより、感情面を含めて強い影響を互いを受ける

ことがある。「人の話を聞くときはちゃんと顔をみなさい」あるいは「ちゃんと相手の目を見て話さない」などと子供の頃誰もが言われた経験があることと思う。このような視線の力を利用して、人とロボットのコミュニケーションをよくしよう、という試みを紹介する。



図1: 視線を制御する Robot

この図のロボットは、NTT-MIT 共同研究の一環で製作されたものであり、多くの自由度と柔軟な制御系、人の顔を検出して追跡する画像処理能力を持っている。この顔追跡機能によって、一見、アイコンタクトが成立したかのような印象をユーザに与えることも可能である。これを使って、アイコンタクトが成立することの効果測定した。測定は、適当な課題(ここでは色名前選択課題)に対して2つの選択肢を与え、ロボットが一方を推奨する。ユーザがこの推奨をどの程度受け入れたか、をロボットの対人間のコミュニケーション能力の一つの指標と考えた。比較実験の結果は、アイコンタクトが成立したときの方が受け入れる率が高まり、効果があることが統計的に確認できた。さらに、人が色板やボタンボックスを見るのにあわせて、同じところを見るようにロボットを動作させ、いわゆる共同注意が成立しやすい条件でも比較した。その結果、共同注意の成立により被験者の判断に強い影響を与えることが、一部の特定の性格の被験者グループに関して確認できた。実際に注意を人に向けること

で、影響力を行使することができ、それは人に限らずロボットやエージェントといったものが行う場合でも同じであることがわかってきた。

注意の応用

次に、映像のハンドリング、特に編集を行う際に、注意情報がどのように役立つかを紹介する。テレビ会議の映像は、せっかく大きな帯域を使って伝送しているにもかかわらず、音声に比べると、人間がそこから受け取れる情報量は限られている。一地点の参加者が多数である場合は特に課題が大きい。画面の中に映る人の顔がとても小さく、表情や反応を読み取ることが困難であると、せっかくのテレビ会議の特長が活かせない。そこで、一人一人をアップで捉え、これを適宜切り替える方法が望ましい。切替には、従来、マイクで音声を捉えて、発話者を映すように切り替える、という手段があったが、十分に会議の展開を反映した切替は実現できていなかった。

ここで我々は、視線多数決に基づく映像切替を提案した。これは、図中にあるように会議参加者の視線を最も集めた人(ここでは、A)にカメラを切り替える、というルールである。すなわち、大勢の注意を集めている人を見つけ、カメラの注意もそこへ向ける、ということになる。これにより、声を発する人のみならず、まさに今説得「され」ようとしている人や、皆が反応を伺いたい、と思っている人へと映像が切り替わり、会議の文脈に合わせ、適切な切替が可能となる。評価実験においても「誰が誰に話しかけていたか」などの情報がより正確に伝わるようになった。

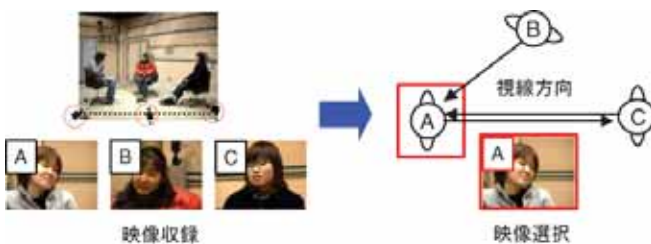


図2：視線に基づく映像切替



(a). 頭部固定版 (b). 頭部非固定版

図3：アイカメラ：FreeGaze

注意の計測

最後に注意の計測手法の研究について述べる。先の視線多数決に基く映像編集を自動的に行うには、会議参加者の視線方向を計測することが必要である。これに対して画像処理を用いて顔の向きを判定する、というアプローチや、より高精度な計測手段として、アイカメラを用いる方法などがある。従来はアイカメラは高精度だがキャリブレーションが煩雑など使いにくい面があった。コミュニケーション科学基礎研究所では、他にはない簡便なキャリブレーション手法を新たに開発し、わずか2点でのキャリブレーションにより十分な精度を実現した。

これにより、視線を用いたインタフェースの実現に一步近づいたものと言える。また、図3(a)では、測定時に頭部位置をある程度固定する必要があったが、図3(b)の頭部非固定版はアイカメラ側が目の位置を追従する機能を持ち、頭の位置を限定しない。

アイカメラは人間の注意が向けられた先を計測・検出するデバイスとして応用範囲は広く、Webの視聴率調査、画面デザインの向上、さらには障害者向け入力装置などとしての期待が高い。

まとめ

人間の注意とその効果を測り、よりよいコミュニケーション実現のために利用する研究は今後の映像コミュニケーション時代の基礎となるものである。今後も着実に成果を重ね、革新的な未来のコミュニケーション実現に貢献したい。

やまと じゅんじ

yamato@eye.br1.ntt.co.jp