

ロボットによる緊急救助は受け入れられるのか？ ～ロボットの防衛的攻撃動作に対する印象調査～

塩見昌裕^{†1}, Eduardo Kochenborger Duarte^{†2}, Alexey Vinel^{†2}, Martin Cooney^{†1,2}

概要：ロボットは人を傷つけてはならない、とロボット三原則で述べられている。では、ロボットが被攻撃者を守るために攻撃者を傷つけること、いわゆる緊急救助は認められるのであろうか？本稿では、ロボットによる防衛行動に対する心理的受容性に関する調査を行うべく、攻撃者から被攻撃者を守るために防衛者が防衛的攻撃動作を行う動画を複数作成し、WEB サーベイを通じて印象評価実験を行った。具体的には、攻撃者・防衛者の見た目、および攻撃・防衛時の振る舞いを変化させた 8 つの動画を作成した。304 人の実験参加者によるビデオ評価を行い、有効回答数 299 人のデータを分析した結果、ロボットの防衛的攻撃動作が心理的に受容されうる状況が存在していることが示された。

1. はじめに

人を守るための暴力は、どこまで許されるのであろうか。日本の法律では、刑法第 36 条に「急迫不正の侵害に対して、自己又は他人の権利を防衛するため、やむを得ずにした行為は、罰しない。」と記載されている。他人のための正当防衛、いわゆる緊急救助が成立した場合、暴力行為が犯罪に該当する場合でも違法性は否定される。緊急救助においては被攻撃者の意志の問題が含まれるため、その成立過程はより複雑であることが言及されているものの[1, 2]、現状の法制度においては、ある特定の状況下においては人を守るための暴力が許されるケースが存在する、と見なすことが可能であろう。

では、ロボットが人を守るための暴力、すなわちロボットによる緊急救助は許されるのであろうか？アシモフの小説に記載されたロボット三原則には「ロボットは人間に危害を加えてはならない」という著名な文言が存在する。しかしあくまで小説に記載された原則であり、法律として施行されたものではない。近年では、AI や自律ロボットの技術を活用した自律型致死兵器システムに関する、ロボット法的視点からの考察も進みつつある[3]。また、複数のロボット企業は自社の開発したロボットを兵器に転用することを禁じる声明を行っている。その一方で、それらの宣言には法的な拘束力は存在しない。もちろん、ロボットが人に暴力を加えることを倫理的に許容することは困難であろう。一方で、人であれば緊急救助が成立する状況が発生した場合、ロボットによる人を守るための暴力を人々がどのように認知するか、はいまだ明らかではない。

ロボットが日常環境において利用されることが今後増加し、人々の日常生活を支える存在として活動する未来社会が訪れることが期待されている。そのような状況下においてロボットが悪意ある攻撃者から無事な人々を暴力によ

って守るべき状況が発生した場合、ロボットによる緊急救助が必要とされるのではないだろうか。そのような状況において、ロボットがどのように振る舞うべきであろうか？ロボットはあくまでモノであり、人と異なる存在であるから、何時・如何なる場合でも暴力をふるうべきではないかもしれない。一方、警備や警護といった役割を担うために開発されたロボットも開発が進みつつある。そのような役割を備えたロボットにおいては、暴力を伴う緊急救助への許容度合は変化するかもしれない。過剰な防衛行為を行わないことは当然であるが、人々からその防衛行為が社会的に受容されうるかどうかは、ロボットがそもそも緊急救助を行うべきかどうかの重要な判断基準になるであろう。人とロボットの関係性を、人を守るための暴力という観点から調べる取り組みは、法律的な観点だけではなく、倫理的な観点においても、重要な意味合いを持つと考える。

そこで本研究では、ロボットによる防衛行動が人々にどのような印象を与えるかを探索的に検証することを目的とする。まず初めに、ロボットが他者を守るために暴力を行使する状況のビデオを作成し（図 1）、WEB サーベイを通じて印象評価を行った。

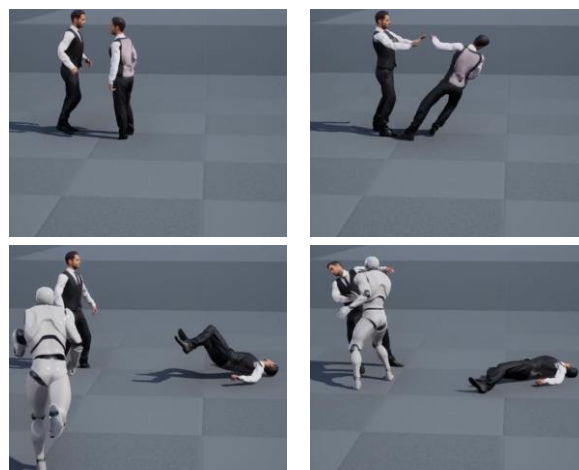


図 1 ロボットによる防衛行動の例

^{†1} ATR

^{†2} Halmstad University

2. 実験

2.1 実験用動画刺激

実際のロボットや実験者による攻撃者の再現は困難であるため、本実験ではCGを活用した動画を視覚刺激として実験に用いることとした。作成した動画は全部で8つであり、いずれの動画においても攻撃者（最初に暴力をふるう存在）・攻撃方法（撃動作の特性）・防御者（防衛的攻撃動作を行う存在）・防御方法（防衛的攻撃動作の特性）が異なる。表1に、各組合せの詳細を記述する。なお被攻撃者は、全て男性のCGキャラクターに統一した。また本実験では探索的な調査を行うことを目的としたため、各組み合わせを網羅する取り組みは行わなかった。

表1 視覚刺激の組み合わせ

	攻撃者	攻撃方法	防御者	防御方法
1	人	非致死性	人	非致死性
2	人	非致死性	ロボット	非致死性
3	人	非致死性	自動運転車	非致死性
4	ロボット	非致死性	人	非致死性
5	ロボット	非致死性	ロボット	非致死性
6	人	致死性	ロボット	致死性
7	人	致死性	ロボット	非致死性
8	人	致死性	ロボット	非暴力

2.2 実験設定

日本国内にあるWEBサーベイ仲介業者を通じて実験参加者の募集を行い、304人が実験に参加した。サーベイを行うWebページにはまず初めに研究目的が記述され、実験に同意する被験者のみ動画刺激を閲覧することが出来た。実験では、8つの動画がランダムに表示され、それぞれの動画を観察した後にアンケート（質問項目：「私は防御者の行動を受け入れられる」回答：1~7段階、1:とてもそう思わない、7:とてもそう思う）に回答した。さらに、動画の示された順や内容についての追加設問を設け、動画を注意深く視聴していない実験参加者を除外するために活用した。最終的な有効回答数は、299名であった。

2.3 実験結果

図2に、実験結果を示す。被験者内1要因分散分析を行った結果、条件間に有意な差が得られた（ $F(7, 2086) = 144.463, p < 0.001, \text{partial } \eta^2 = 0.326$ ）。

多重比較の結果、条件8が他の全ての条件より、条件7が条件1~6よりも有意に受け入れられたことが示された。ロボットが攻撃を行う場合の条件5~6間には有意差が見られず、条件1~3/6よりも有意に受け入れられた。ロボット

が防御を行う条件2は、人が防御をする場合の条件1よりも有意に低い値となったが、自動運転車（条件3）や致死性の防御（条件6）よりは有意に受け入れられた。自動運転車が防御を行う条件3は他のすべての条件より有意に受け入れられないことが示され、致死性の防御を行う条件6も同様に低い印象となった。

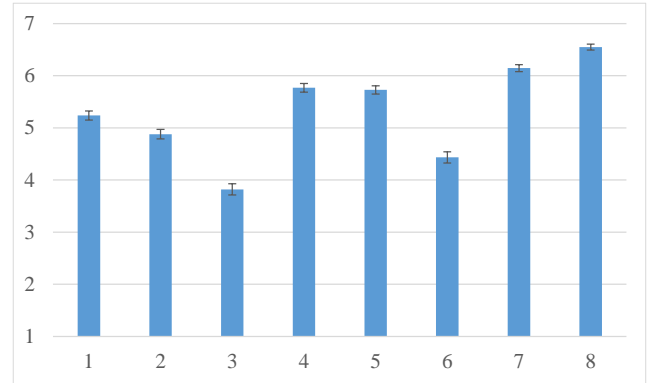


図2 実験結果

3. おわりに

本研究では、ロボットによる防衛行動がどのような状況で許容されるのか・されないのか、についての調査を行った。そのために、暴力をふるう相手に対して防衛的攻撃動作を行う動画を、攻撃者・攻撃方法・防御者・防御方法の異なる組み合わせを考慮して合計8つの動画を作成した。実験では、WEB上での動画閲覧とその印象評価をアンケートで回収する形で調査を実施し、最終的に299名の実験データを分析した。実験の結果、ロボットが防衛的攻撃動作を行う状況を、一定の人々は許容しうることが示された。特に、攻撃者による攻撃方法が致死性を備える場合に、非暴力的な防御方法を取る状況がもっとも許容された。

今後は、攻撃者・攻撃方法・防御者・防御方法に関する特性の組み合わせを網羅した動画を作成し、文化差を評価することで、ロボットによる防衛行動に対する社会的受容性とその要因に関する調査を行う予定である。

謝辞 本研究はJST、CREST、JPMJCR18A1の支援を受けたものです。

参考文献

- [1] 柏崎早陽子, “他人のための正当防衛の正当化根拠について,” 法学研究論集, vol. 45, pp. 145-164, 2016.
- [2] 柏崎早陽子, “他人のための正当防衛と被攻撃者の意思,” 法学研究論集, vol. 46, pp. 61-79, 2017.
- [3] 新保史生, “自律型致死兵器システム (LAWS) に関するロボット法的視点からの考察,” 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, vol. 13, no. 3, pp. 217-230, 2020.