様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ２９ 年　 ８ 月 ２６ 日現在

機関番号：３２６７８
研究種目：基盤研究(C)（一般）
研究期間：2013～2016
課題番号：２５３３０２０６

研究課題名（和文）模倣音声による詐称攻撃に対して頑健な話者照合の研究

研究課題名（英文）Research on robust speaker verification against spoofing attacks by voice imitation

研究代表者
岩野　公司（Iwano, Koji）

東京都市大学・メディア情報学部・教授

研究者番号：９０３２３８２３

交付決定額（研究期間全体）：（直接経費）　　3,400,000円

研究成果の概要（和文）：本研究では，話者照合システムに対する声真似（模倣音声）による成りすまし（詐称）攻撃の影響の分析と，模倣音声の音響特徴の分析を行った．本研究で独自に収録した一般人とプロの物真似タレントの模倣音声データを用いて分析を行った結果，一般人の模倣でも成りすましの成功確率が有意に上昇し，その影響が無視できないこと，プロのタレントの模倣は一般人よりも攻撃力が大きいことがわかった．また，「模倣のうまさ」を定量的に評価する手法を提案し，それによってプロの声真似が効率的に対象者の声質に近づいていることを明らかにした。

研究成果の概要（英文）：In this research, we experimentally analyzed the effect of spoofing attacks by voice imitation on speaker verification systems and acoustical characteristics of the imitated voice. These analyses were conducted by using our original speech data consisting of professional and non-professional impersonators' imitated voice. The analysis results show that the voice imitation by non-professional impersonators significantly increases the success rate of spoofing attacks and the higher success rate is yielded by the professional impersonator's imitation. We also proposed a method for quantitatively evaluating the quality of voice imitation. The method reveals that the professional imitator efficiently approaches his voice characteristics towards the target speaker's voice.

研究分野：音声情報処理

キーワード：話者照合　模倣音声　詐称攻撃　話者認識　個人認証

１．研究開始当初の背景

近年，生体情報を利用した個人認証は，認証子の紛失や忘却の危険性が少ないため，セキュリティシステムへの積極的な導入や実用化が進められている．その中でも「声による個人認証（話者照合）」は，特殊な入力装置が不要で，安価・簡易な導入が可能であることから，様々な情報システムの手軽な認証方式として注目され，世界的に研究が進められている．

話者照合システムの構築と性能評価は，多数の話者による，複数時期で発声された音声データを利用して行われる．従来の話者照合用の音声データベースでは，各話者の音声は「システムに受理されるべき申告者本人の声」として収録されている．一方，話者照合性能を評価する際には，「申告者本人の発声が他人として棄却される（本人棄却）誤り」だけでなく「他人が申告者に成りすまして（詐称して）行った発声が申告者として受理される（詐称者受理）誤り」を考慮する必要がある．しかし，これまでの話者照合の研究では，それぞれの話者が申告者本人の声として発声したデータを，他人が申告者であったときの詐称者発声として便宜的に使用しているため，詐称者受理誤りを「詐称・偽造の意図を持っていないデータ」で評価している，という問題があった．

意図を持った詐称に対する生体認証性能の検証は非常に重要な課題であり，例えば，音声については，音声合成器を利用した詐称に対する脆弱性の指摘とその対策手法が検討されている．しかし，もっとも単純な詐称手段である「模倣（声真似・声帯模写）」という攻撃に対する照合の脆弱性についてはほとんど研究が進められていない．

２．研究の目的

本研究では，「本人としてシステムに受理されようとする発声」の他に，「声真似によって他人に成りすまし，システムを攻撃しようとする発声（模倣音声）」を収録したデータベースの構築を行い，それを用いて模倣攻撃によって照合性能の劣化がどの程度生じるかを明らかにする．

また，模倣発声のモデル化と分析を行い，模倣によって話者内の音響特徴量空間がどの程度変動するか，「模倣のうまさ」と特徴量空間の変動度合いの関係性（図１）を明らかにし，模倣音声による攻撃への対策について考察する．

３．研究の方法

(1) まず，通常の発声（地声）と模倣音声で構成される音声データベースの構築を行い，実験基盤の整備を行う．一般性の高い知見を得るため，模倣音声として様々なレベルの音声を取り扱う．具体的には，「声真似に特殊な技能を有していない者（一般人）による模倣」，「一般人が模倣対象者への声真似の訓練

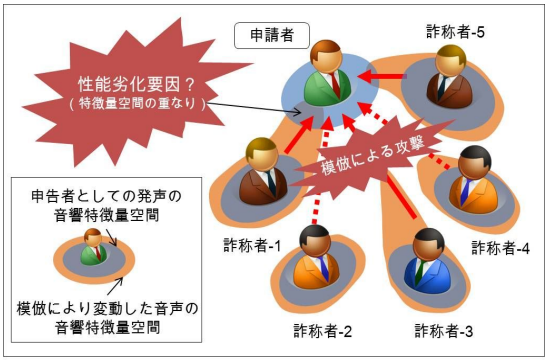

図１：模倣による音響特徴量空間の変動

を行った場合の模倣」，「日ごろから声真似の訓練を行っているプロの物真似タレントによる模倣」の３種類の音声収録を行った．それぞれについて，模倣の対象者に共通の人物を設定することで，３者の比較を厳密に行うことができる．

(2) (1) で構築したデータベースを用いて，GMM-UBM 法による話者照合システムの構築し，模倣音声による詐称攻撃が話者照合性能に与える影響の評価を行う．当初の実験計画では i-vector に基づく照合システムでの評価も実施予定であったが，予備実験の結果，入力音声の発話長が短いことに起因して，十分なベースラインの性能が得られなかったことから，GMM-UBM 法のみでの評価に変更した．

(3) (2) で得られた結果，特に，一般人とプロの模倣の「うまさ」の違いが，どのような現象によって引き起こされているかを明確にするため，模倣に伴う音声特徴量の変動を定量的に分析する手法を提案し，(1) で収録したデータを用いて分析を行うことで知見を得る．その知見をもとに，模倣による攻撃の対策について考察を行う．

４．研究成果

(1) 話者照合用データベースとして，

① 女子大学生（一般女性）６名の地声と，自分以外の５名への模倣音声
② 男子大学生（一般男性）６名の地声と，自分以外の５名への模倣音声
③ ①と同じ男性６名による，地声と自分以外の５名への模倣音声（再収録），さらに訓練を行った上での模倣音声
④ プロの男性物真似タレント１名による，①の男性６名への模倣音声

を収録し，データベースの構築を行った．発声内容は４桁の連続数字である．

①～③については，約２日ごと約３週間にわたって（合計９日）収録を行った．各日について，地声と各話者への模倣音声をそれぞれ１０発声ずつ行っている．なお，それぞれの被験者は声質の近さなどは考慮せずに選ばれている．④については，１日のみでの収録となっており，地声と各話者への模倣音声をそれぞれ１０発声ずつ行っている．プロは

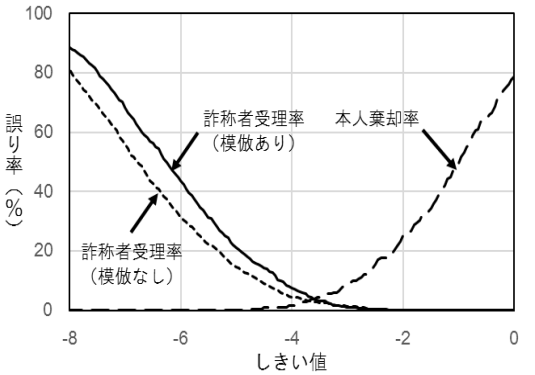

図２：一般女性６名の模倣（データ①）による詐称者受理率の様子

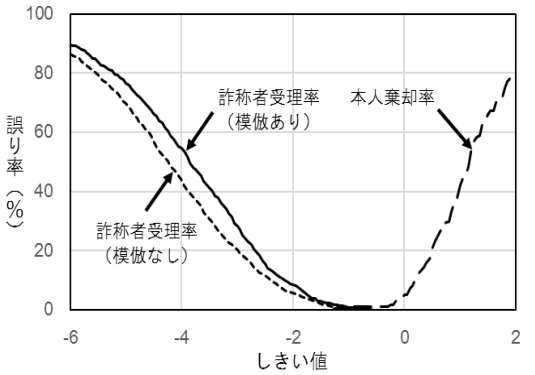

図３：一般男性６名の模倣（データ②）による詐称者受理率の様子

事前に模倣の対象者とは面識がなく，収録の場で音声を聴取しながら模倣を行っている．なお，①，②のデータについては複数人の被験者による聴取による「模倣のうまさ」のスコアが付与されている．

③における模倣の訓練では，前日までに発声された地声のデータで構築された話者照合システムを利用し，そのシステムから得られる照合スコアが高くなるように（一時的に）発声の工夫を見出した上で，模倣音声の収録を行った．

構築したデータベースは，プロのレパートリーではない話者に対し，一般人とプロが共通に，対等な条件で模倣を行っている点，一般人が物真似の訓練を行った際のデータが収録されている点に独自性があり，これまでに例が無い貴重なデータとなっている．

(2) (1)で収録したデータについて，①～③の前半の複数日で収録されたデータを用いて GMM-UBM 法による話者照合システムを構築し，①～④の（システム構築に用いていない）データを利用して照合性能の評価を行った．実験の結果，

① 一般人の模倣攻撃であっても，模倣攻撃を想定していない場合に比べ，成りすましの成功確率（詐称者受理率）が数%上昇することが確認され，実用上無視できない影響を与えること（図２，３）．

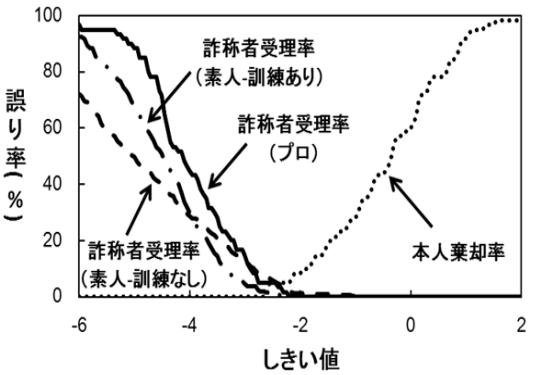

図４：一般男性６名の訓練あり／なしの模倣（データ③）と，プロの物真似タレントの模倣（データ④）による詐称者受理率の比較

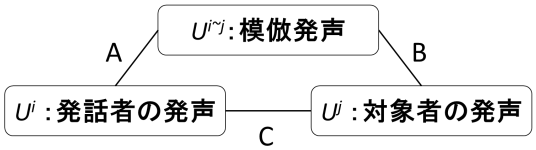

図５：分析対象とする発声間距離（A～C）

② 一般人であっても訓練を行うことで模倣の攻撃力が増し，訓練を行わない模倣に比べて詐称者受理率の上昇が見られること（図４）．

③ プロの模倣音声は，一般人の模倣（訓練あり／なし）よりも大きな詐称者受理率の向上を招くこと（図４）．

が確認された．

一般人による模倣の攻撃力を明らかにした点，物真似の対象者を共通化して，一般人とプロの攻撃力を厳密に比較した点が，本研究の独自の成果である．

(3)模倣時の音声特徴量（話者照合で用いられるケプストラム特徴量）の変動を分析するため，「カルバック・ライブラー情報量を用いた発声間距離に基づく分析手法」を提案した．この手法では，「発話者 i の発声（地声）」「対象者 j の発声（地声）」「発話者 i が対象者 j を真似したときの模倣発声」の３つの発声に対して音響モデルを構築し，それぞれのモデル間の距離を調べる手法である．３つの距離の関係を図５に示す．

発声間距離の算出では，まず，分析対象音声をケプストラムに基づく音響特徴量（12 次元 MFCC＋12 次元ΔMFCC＋Δ対数パワー）に変換し，各発声を３状態の隠れマルコフモデルでモデル化する．その２状態目の混合正規分布間の距離を，対称化カルバック・ライブラー情報量を利用して算出し，発声間距離とする．(1)の③，④で収録された男性の音声データを用いて発声間距離を算出した結果，

① 一般人は声真似によって自身の声を大きく変動させることはできているが，対象者の声質には近づいていないこと．

② プロの物真似タレントは一般人ほど声質を大きくは変化させていないが，対象者には確実に近づいていること．

が明らかになった．この結果は(2)で得られた話者照合システムの詐称者受理率の傾向とも矛盾がないことがわかる．

これは，「話者照合システムをだます」という観点での「模倣のうまさ」を，提案する分析手法によって（事前に話者照合システムを構築する必要なく）定量的な表現として数値化できることを示唆している．

以上をまとめると，(2)の結果より，話者照合システムにおいては，一般人の模倣についてもその攻撃への対策が必要であり，(3)の結果から一般人の模倣については発声内，発声間での特徴量変動が大きいことが示唆されることから，今後はその特徴を利用した模倣攻撃への対策手法の考案を行う必要がある．一方で，プロの模倣についての対応が上記のような対策で十分であるかについては不明確であり，今後さらに検討を重ねていく必要がある．

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計　０　件）

〔学会発表〕（計　５　件）

① 堀畑拓斗，岩野公司，“話者照合におけるプロの物真似タレントの声真似攻撃の影響の分析，”情報処理学会第 79 回全国大会，2017 年 3 月 18 日，名古屋大学（愛知県・名古屋市）．
② Koji Iwano，Takuto Horihata，“Analysis of Voice Imitation by Professional/Non-Professional Impersonators Based on Kullback-Leibler Divergence between Acoustic Models，” the 5th Joint Meeting of Acoustical Society of America and Acoustic Society of Japan (ASA/ASJ)，2016 年 11 月 29 日，ホノルル（アメリカ合衆国）．
③ 曽根泰斗，岩野公司，“声真似攻撃に対する話者照合システムの脆弱性の分析，”電子情報通信学会総合大会，2015 年 3 月 11 日，立命館大学（滋賀県・草津市）．
④ 岩野公司，曽根泰斗，坂本香菜子，“声真似が話者照合に与える影響と物真似音声の音響特徴の分析，”電子情報通信学会音声研究会，2015 年 1 月 22 日，じゅうろくプラザ（岐阜県・岐阜市）．
⑤ 坂本香菜子，岩野公司，“話者照合への影響を考慮した模倣音声の音響分析，”情報処理学会第 76 回全国大会，pp.473-474，2014 年 3 月 13 日，東京電機大学（東京都・足立区）．

〔図書〕（計　０　件）

〔産業財産権〕
○出願状況（計　０　件）

○取得状況（計　０　件）

〔その他〕
ホームページ等
http://www.yc.tcu.ac.jp/~iwanolab/

６．研究組織
(1)研究代表者
岩野　公司（IWANO，Koji）
東京都市大学・メディア情報学部・教授
研究者番号：９０３２３８２３

(2)研究分担者
篠田　浩一（SHINODA，Koichi）
東京工業大学・情報理工学院・教授
研究者番号：１０３４３０９７