音声対話システムによる精神疾患自動評価

Automated Psychiatric Evaluation through Spoken Dialogue System

横谷　謙次(Yokotani Kenji)[1]

[1] 新潟青陵大学大学院臨床心理学科(Niigata Seiryo University)

Abstract

Daily conversation is used as a main model of Spoken Dialogue System. The present study aims to show that psychiatric conversation can be a model of the system. We implemented Mini International Neuropsychiatric Interview in the system and suggest that automated psychiatric evaluation through the system can be linked with professional psychiatric evaluation with the structured clinical interview for DSM-4 (SCID 2). Application of the system into psychiatric fields could extend the system's user base.

1 はじめに

音声対話システムの研究は今後、キーボードやマウスよりも優位性があり、かつ、日常的に使用されるアプリケーションが開発され、そして、そのアプリケーションがビジネスモデルとして成立することが期待されている[1]。

本研究で開発した精神疾患自動評価の音声対話システム（以下、自動評価システム）は、画像・音声情報を含んでいるため、キーボードやマウスよりも多様な情報を含んでおり、優位性がある。また、日本では、勤労者へのストレスチェックが平成２５年度より義務付けられているため、そのストレスチェックの一貫として同システムが使用されれば、日常的なアプリケーションになり得る。加えて、今まで専門家が行っていた精神疾患の評価を同システムが同程度の精度で評価し得ることを示せば、人件費のかからない同システムを利用する企業が出てくる可能性があり、ビジネスモデルとしても成立するだろう。

2 精神疾患自動評価システム

コンピューターによる自動診断は 1960 年代から開発されてきた[2]。1980 年代から現在にいたるまでのほとんどのアプリケーションは放射線や顕微鏡や CT スキャンなどの画像データを専門家が取得し、その取得された画像をコンピューターが自動解析するという受動的なものであった[3] [4]。

図 1.MINI を組み込んだ精神疾患自動評価システム

本研究では、音声対話システムを用いて、対象者の画像、音声及び言語データをコンピューターが自ら取得し、自動で解析する（図１参照）。本研究は、コンピューターが自らデータを取得する点でこれまでのアプリケーションよりも能動的であり、コンピューターの役割を拡張させたと言える。特に表情データを数十ミリ秒単位で取得する

ⓒ 2016 Information Processing Society of Japan

ことにより、表情で示される情動変化を微細に測定することが出来る[5]–[7]。また、音声データの取得によっても、音声の観点から情動を測定することも出来る[8], [9]。もちろん、言語データからも情動を測定することが出来る[10]。対象者の表情、音声、及び言語が精神疾患と関連することは古くから指摘されている[11], [12]。

自動評価システムの質問項目として、Mini International Neuropsychiatric Interview (M.I.N.I. 5.00)の日本語版の一部を使用した[13], [14]。MINI は「はい」「いいえ」のどちらかで答えられる質問でほとんど構成されており、15 分ほどの対話で終了する。また、対象者の気分障害、不安障害、アルコール依存症、摂食障害を評価することが出来、これらの精神疾患は日本の勤労者がかかり易い精神疾患をほとんど網羅している[15]。

また、自動評価システムの評価の妥当性を確認するために、専門家による精神疾患評価を同時に行う。Structured clinical interview for DSM-IV(SCID4)日本語版を用いて[16], [17]対象者の精神疾患を臨床心理士が評価する。SCID4 は世界中で使用されている構造化面接法である。

本研究には三つの目的がある。1.自動評価システムの評価と専門家による評価は関連する。2.同システム及び専門家による評価中の対象者の表情は精神疾患と関連する。3.同様に対象者の音声及び自然言語も精神疾患と関連する。

本研究では、大学生が自動評価システムによる評価と専門家による評価を受ける。大学生のうち半数は精神疾患自動評価システムを受けたあと、専門家による評価を受ける。別の半数は専門家の評価を受けた後に自動評価システムを受ける。両評価中は常に録音・録画されている。

なお、同システムの開発環境は unity であり、音声認識として Julius の version4.4 を用いた[18]。現在、本研究は倫理審査中である。

3 解析方法

3-1.自動評価システムによる評価と専門家による評価との関連性

自動評価システムと専門家の評価とを比較する。大うつ病性障害を例にすれば、表 1 を用い、SCID 及び MINI との関連を見る。自動評価システムの 18 個の精神疾患にこの表を適用し、精神疾患毎の精度を確認する。

表 1.大うつ病性障害の比較

		専門家による SCID	
自動評価		－	＋
による	－	a	b
MINI	＋	c	d

3-2. 表情認識

20fps で面接中(精神疾患自動評価システム及び専門家によるインタビュー)の写真を撮影し、表情評価を行う（図 2 参照)。表情評価には Microsoft Cognitive Service の Emotion API を用いる[19]。この方法によって、約 50 ミリ秒毎の対象者の表情を客観的に評価することが出来る。

この測定手法は既存の手法よりも対象者への侵襲性が低く、かつ、観察単位も細かいため、対象者の情動をより微細に測定することが出来る（表 2 参照)。

3-3. 音声

面接中の音声を録音し、対象者の基本周

波数や声門破裂音の観点から解析する[20]。抑うつと評価された群とそうでない群との間の違いを検討する。

3-4.自然言語

また、面接中の言語を収集し、抑うつ群とそうでない群との違いを検討する。人称代名詞の使用に着目し[12]、SCID における自由報告場面の内容を検討する。

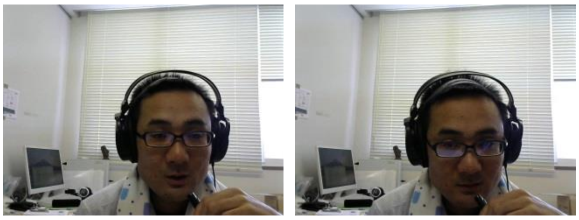

図 2.MINII 実施中の表情

表 2.図 2 の表情に基づく情動評価

	図 2 左	図 2 右
Anger	0.001892032	0.001094373
Contempt	0.004701399	0.002108159
Disgust	0.001176891	0.0003307918
Fear	0.0002366242	0.0001572938
Happiness	0.001356887	0.0007938205
Neutral	0.8549111	0.9470335
Sadness	0.002650511	0.00087655
Surprise	0.1330745	0.04760551

4　おわりに

精神疾患自動評価システムは、勤労者の精神保健領域をカバーしており、精神保健のストレスチェックの補助として使用しうる。自動評価システムは専門家による評価よりも人件費がかからないため、同システムを勤労者の精神保健に適用することは可能だろう。また、同システムの使用は、勤労者の音声対話システムの日常的な使用を促すことにも役立ちうる。

謝辞

本研究は科学技術融合振興財団の研究支援を受けております。

参考文献

[1] 達也　河原(kawahara), “音声対話システムの進化と淘汰：歴史と最近の技術動向(<特集>音声対話システムの実用化に向けて),” 人工知能学会誌, vol. 28, no. 1, pp. 45–51, Jan. 2013.

[2] K. Doi, “Computer-aided diagnosis in medical imaging: historical review, current status and future potential,” *Comput. Med. Imaging Graph.*, vol. 31, no. 4, pp. 198–211, 2007.

[3] N. Sharma and L. M. Aggarwal, “Automated medical image segmentation techniques,” *J. Med. Phys.*, vol. 35, no. 1, p. 3, Jan. 2010.

[4] D. Lu, X.-H. Yu, X. Jin, B. Li, Q. Chen, and J. Zhu, “Neural network based edge detection for automated medical diagnosis,” in *Information and Automation (ICIA), 2011 IEEE International Conference on*, 2011, pp. 343–348.

[5] P. Ekman, “Darwin, deception, and facial expression,” *Ann. N. Y. Acad. Sci.*, vol. 1000, no. 1, pp. 205–221, 2003.

[6] P. Ekman, D. Matsumoto, and W. V. Friesen, “Facial expression in affective disorders,” *What Face Reveals Basic Appl. Stud. Spontaneous Expr. Using Facial*

Action Coding Syst. FACS, vol. 2, pp. 331–342, 1997.

[7] S. Polikovsky, Y. Kameda, and Y. Ohta, "Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor," in *Crime Detection and Prevention (ICDP 2009), 3rd International Conference on*, 2009, pp. 1–6.

[8] A. Ozdas, R. G. Shiavi, S. E. Silverman, M. K. Silverman, and D. M. Wilkes, "Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1530–1540, Sep. 2004.

[9] J.-A. Bachorowski and M. J. Owren, "Vocal expressions of emotion," *Handb. Emot.*, vol. 3, pp. 196–210, 2008.

[10] S. Rude, E.-M. Gortner, and J. Pennebaker, "Language use of depressed and depression-vulnerable college students," *Cogn. Emot.*, vol. 18, no. 8, pp. 1121–1133, 2004.

[11] F. Schneider, H. Heimann, W. Himer, D. Huss, R. Mattes, and B. Adam, "Computer-based analysis of facial action in schizophrenic and depressed patients," *Eur. Arch. Psychiatry Clin. Neurosci.*, vol. 240, no. 2, pp. 67–76, 1990.

[12] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, "Psychological aspects of natural language use: Our words, our selves," *Annu. Rev. Psychol.*, vol. 54, no. 1, pp. 547–577, 2003.

[13] D. V. Sheehan *et al.*, "The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability," *Eur. Psychiatry*, vol. 12, no. 5, pp. 232–241, 1997.

[14] T. Otsubo *et al.*, "Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview," *Psychiatry Clin. Neurosci.*, vol. 59, no. 5, pp. 517–526, Oct. 2005.

[15] 川上憲人, "勤労者における精神疾患の疫学：頻度と仕事関連要因 (特集 勤労者のメンタルヘルス)," *日本医師会雑誌 J. Jpn. Med. Assoc.*, vol. 144, no. 12, pp. 2437–2441, Mar. 2016.

[16] M. B. First, R. L. Spitzer, M. Gibbon, and J. B. Williams, *User's guide for the Structured clinical interview for DSM-IV axis I disorders SCID-I: clinician version*. American Psychiatric Pub, 1997.

[17] 松岡奈緒 and 北村俊則, "分子精神医学のための臨床評価 (1) Structured Clinical Interview for DSM-4 (SCID)," 分子精神医学, vol. 5, no. 1, pp. 47–50, 2005.

[18] "大語彙連続音声認識エンジン Julius." [Online]. Available: http://julius.osdn.jp/. [Accessed: 15-Nov-2016].

[19] "Microsoft Cognitive Services - Emotion API." [Online]. Available:

https://www.microsoft.com/cognitive-services/en-us/emotion-api. [Accessed: 15-Nov-2016].

[20] E. Moore II, M. A. Clements, J. W. Peifer, and L. Weisser, "Critical analysis of the impact of glottal features in the classification of clinical depression in speech," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 96–107, 2008.

ⓒ 2016 Information Processing Society of Japan