Daftar Pustaka

[1] Youtube, *YouTube for Press*, [Online] Available At: https://www.youtube.com/yt/press/en-GB/statistics.html [Accessed 22 March 2017]

[2] Robertson, M.R., 2015, *500 Hours of Video Uploaded to Youtube Every Minute [Forecast]*, [Online] Available At: http://tubularinsights.com/hours-minute-uploaded-youtube/ [Accessed 22 March 2017]

[3] Cisco, 2017, *Cisco Visual Networking Index: Forecast and Methodology, 2016–2021*, White paper Cisco public

[4] Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., Povey, D. and Valtchev, V., 2002. *The HTK book*. Cambridge university engineering department, 3, p.175.

[5] Janakiraman, R., Kumar, J. C., Murthy, H. A., 2010, *Robust syllable segmentation and its application to syllable-centric continuous speech recognition*, Chennai, National Conference On Communications (NCC).

[6] Mak, M.W., Allen, W.G., 1994, *Lip-motion analysis for speech segmentation in noise*, Speech Communication, Volume 14, Issue 3, Pages 279-296, ISSN 0167-6393.

[7] Talea, H., Yaghmaie, K., 2011, *Automatic visual speech segmentation*, 2011 IEEE 3rd International Conference on Communication Software and Networks, Xi'an, pp. 184-188.

[8] Muda, L., Begam, M., Elamvazuthi, I., 2010, *Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques*, Journal of Computing, Volume 2, Issue 3, ISSN 2151-9617.

[9] Han, W., Chan, C.F., Choy, C.S., Pun, K.P., 2006, *An Efficient MFCC Extraction Method in Speech Recognition*, Island of Kos, 2006 IEEE International Symposium on Circuits and Systems

[10] Devcore, 2012, *8x8 DCT (discrete cosine transformation)*, [Online] Available at: https://en.wikipedia.org/wiki/File:DCT-8x8.png [Accesed 26 August 2018]

[11] Wallace, G.K.,1992, *The JPEG Still Picture Compression Standard*, Massachusetts, IEEE Transactions on Consumer Electronics, Vol. 38, No. 1.

[12] El Maghraby, E.E., Gody, A.M., Farouk, M.H., 2016, *Enhancing Quality and Accuracy of Speech Recognition System By Using Multimodal Audio-Visual Speech Signal*, Cairo, 2016 12th International Computer Engineering Conference (ICENCO), pp. 219-229.

[13] Adjeroh, D.A., Lee, M.C., 2004, *Scene-Adaptive Transform Domain Video Partitioning*, IEEE Transactions on Multimedia, vol. 6, no. 1, pp. 58-69.

[14] Sharif, M., Mohsin, S., Jamal, M.J., Raza, M., 2010, *Illumination Normalization Preprocessing for Face Recognition*, Wuhan, 2010 The 2nd Conference on Environmental Science and Information Application Technology, pp. 44-47.

[15] Wang, B., Li, W., Yang, W., Liao, Q., 2011, *Illumination Normalization Based on Weber's Law With Application to Face Recognition*, IEEE Signal Processing Letters, vol. 18, no. 8, pp. 462-465.

[16] Wold, S., Esbensen, K., Geladi, P., 1987, *Principal component analysis*, Sweden, Chemometrics and Intelligent Laboratory Systems, Volume 2, Issues 1–3, Pages 37-52.

[17] Smith, L., 2002, *A Tutorial On Principal Components Analysis*, USA, Cornell University.