Appendix A

Database

Yale Face Database

Figure A.1: One individual with 11 different facial variations from Yale face database is shown.

The Yale face database contains 165 gray scale images of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, with-glasses, happy, left-light, with-no-glasses, normal, right-light, sad, sleepy, surprised, and wink. Fig.(A.1) shows 11 images of one individual from Yale database.

\footnote{cvc.yale.edu/projects/yalefaces/yalefaces.html}
Advanced Multimedia Processing (AMP) Facial Expression Database\textsuperscript{1}

The database consists of 13 subjects, whose facial images were captured with varying expressions. Each subject in the database has 75 images of varying facial expressions. The faces were captured in a video sequence where a face tracker\textsuperscript{169} tracked the movement of the user’s head and based upon an eye localization routine and extracted registered face images of size $64 \times 64$. Example images of one person are shown in Fig.(A.2).

\textbf{Extended Yale Face Database B}\textsuperscript{2}

The Extended Yale Face Database B\textsuperscript{142} contains 38 individuals under 64 different illumination conditions with 9 poses. Only frontal face images are taken for experiments. Frontal cropped face images are readily available from website. The cropped faces are categorized in five different subsets according to the angle between the light source direction and the camera axis. Fig.(A.3) shows five subsets of 64 different faces of one individual from Extended YaleB database.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{example_images.png}
\caption{Example images of one individual from AMP facial expression database are shown with drastic expression variations.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{five_subsets.png}
\caption{Five subsets of 64 different faces of one individual from Extended YaleB database.}
\end{figure}

\textsuperscript{1}\url{amp.ece.cmu.edu/projects/FaceAuthentication/download.htm}
\textsuperscript{2}\url{http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html}
Figure A.3: One individual under 64 different lighting conditions from YaleB database is shown. The angle between light source direction and camera axis is given at the top of corresponding subset.

Pose Illumination Expression Face Database

The PIE database[143] contains 41,368 images of 68 people, each person under 13 different poses, 43 different illumination conditions, and with 4 different expressions. A subset of this database is used here for experimental purpose. This subset contains 65 people each with 21 different illumination conditions with frontal face posing and neutral expression with room lights off. Another set of similar dataset can be used with room lights on. The dataset with room lights off is comparatively harder dataset and used for experimental purposes in this study. The original $640 \times 486$ pixel color images are converted to gray scale, since only the intensity information is considered. Fig.(A.4) shows one person from PIE database having 21 different lighting conditions with neutral expression.

\footnote{vasc.ri.cmu.edu/idb/html/face/}
Figure A.4: One individual from PIE database with 21 different illumination condition under room light off is shown.

AR Face Database\(^1\)

The AR face database[156] contains over 4000 face images of 126 people, including frontal view of faces with different facial expressions, lighting conditions and occlusions. The images of 120 individuals were taken in two sessions (separated by two weeks) and each session contains 13 color images. 14 face images (each session contains seven images) are selected from each of these 120 individuals. Fig.(A.5) shows sample images corresponding to one session of two persons.

Figure A.5: One individual from AR database is shown with different lighting and drastic expression variations.

\(^1\)http://www2.ece.ohio-state.edu/aleix/ARdatabase.html
VidTIMIT Audio-Video Database

The VidTIMIT dataset[167] is comprised of video of 43 people, reciting short sentences. There are 10 sentences per person. The first six sentences are assigned to Session 1. The next two sentences are assigned to Session 2 with the remaining two to Session 3. The first two sentences for all persons are the same, with the remaining eight generally different for each person. This data set is useful for automatic lip reading, small movement of facial parts, multi-view face recognition and person identification. Fig.(A.6) shows some frames of one person correspond to one video of VidTIMIT audio-video database.

Figure A.6: Some frames of one individual from VidTIMIT database are shown with different facial expression variations.

1http://itee.uq.edu.au/conrad/vidtimit/