

ÍNDICE DE FIGURAS

<i>Figura 1. Modos de funcionamiento de un sistema de reconocimiento automático de locutor. a) Identificación y b) Verificación. Figura adaptada de [D.A. Reynolds, 2008]</i>	15
<i>Figura 2. Diagrama de bloques de la fase de entrenamiento de un sistema de verificación automática de locutor. Figura adaptada de [Frédéric Bimbot et al., 2004]</i>	16
<i>Figura 3. Diagrama de bloques de la fase de test de un sistema de verificación automática de locutor. Figura adaptada de [Frédéric et al., 2004]</i>	17
<i>Figura 4. Ejemplo para ilustrar las dos fases de un sistema de verificación automática de locutor</i>	17
<i>Figura 5. Curvas de densidad de probabilidad de las puntuaciones de usuarios e impostores</i>	19
<i>Figura 6. Curvas ROC y DET</i>	20
<i>Figura 7 Ejemplo de una curva DET con los correspondientes valores de EER y minDCF</i>	21
<i>Figura 8. Diagrama de bloques de la parametrización mediante MFCC. Figura adaptada de [Eugenio Arévalo, 2011]</i>	24
<i>Figura 9. Banco de filtros utilizado por [S. B. David and P. Mermelstein, 1980]</i>	26
<i>Figura 10. Escala de frecuencias MEL</i>	27
<i>Figura 11. Banco de filtros MEL</i>	27
<i>Figura 12. Los dos ejemplos de actividad de detección de energía (SAD): weightSAD y meanSAD</i>	30
<i>Figura 13. Distribución de la primera componente cepstral de 2.5 minutos de conversación de un locutor antes de ser sometida a SAD</i>	32
<i>Figura 14. Distribución de la primera componente cepstral de 2.5 minutos de conversación después de ser sometida a SAD y normalización mediante CMVN</i>	33
<i>Figura 15. Modelo de mezcla gaussiana mono-dimensional para una función de distribución de entrada. Histograma (a la izquierda) junto con dos aproximaciones GMM de diferente tamaño (4 y 32, respectivamente) [Robin Harald Priewald, 2009]</i>	37

<i>Figura 16. Modelo de mezclas gaussianas bidimensional para una función de distribución de entrada. Histograma (a la izquierda) junto con dos aproximaciones GMM de diferente tamaño (centro y derecha) [Robin Harald Priewald, 2009]</i>	<i>37</i>
<i>Figura 17. Ilustración sobre los tres tipos de segmentos o grabaciones necesarias en GMM-UBM.....</i>	<i>38</i>
<i>Figura 18. En rojo, generación del modelo universal en GMM, y en azul, la fase de entrenamiento para un locutor específico mediante adaptación de ciertos parámetros del UBM. Figura adaptada de [B. Fauve, 2009].</i>	<i>39</i>
<i>Figura 19. Ejemplo de adaptación de un modelo de locutor. Adaptada de [D.A. Reynolds, T.F. Quatieri and R.Dunn, 2000]</i>	<i>42</i>
<i>Figura 20. Ilustración esquemática de la fase de test en la técnica de GMM-UBM. Figura adaptada de [B. Fauve, 2009].....</i>	<i>44</i>
<i>Figura 21. Posibles hiperplanos de separación para puntos bidimensionales pertenecientes a dos tipos de clase.....</i>	<i>46</i>
<i>Figura 22. Hiperplano óptimo para el caso representado en la figura 20.....</i>	<i>46</i>
<i>Figura 23. Transformación de un espacio bidimensional a otro tridimensional a través de la función interna de un Kernel. Figura adaptada de [Alejandro Abejón, 2007].....</i>	<i>47</i>
<i>Figura 24. Ejemplo de puntuación en máquinas de vectores soporte.....</i>	<i>48</i>
<i>Figura 25. Ilustración de la obtención del GMM-supervector. Figura obtenida de [Campbell et al., 2006c].....</i>	<i>50</i>
<i>Figura 26. Obtención de los supervectores de entrenamiento (azul) y de test (verde) en SVM-GMM.</i>	<i>50</i>
<i>Figura 27. Principios generales de SVM aplicados a Verificación Automática de Locutor. Figura adaptada de [B. Fauve, 2009].</i>	<i>51</i>
<i>Figura 28. Ejemplo de NAP con distintas grabaciones en diferentes condiciones para 3 locutores.</i>	<i>53</i>
<i>Figura 29. Resta de todas las locuciones de cada usuario por su correspondiente supervector media en NAP. Figura adaptada de [B. Fauve, 2009].....</i>	<i>54</i>
<i>Figura 30. Proyección en NAP para un ejemplo de supervectores bidimensionales. ...</i>	<i>55</i>
<i>Figura 31. Estructura de la normalización Z_{norm}.</i>	<i>60</i>
<i>Figura 32. Estructura de la normalización T_{norm}.</i>	<i>61</i>

<i>Figura 33. Porción del fichero 1conv4w.trn correspondiente a la carpeta train de la evaluación del NIST de 2005.....</i>	<i>64</i>
<i>Figura 34. Porción del fichero 1conv4w-1conv4w.ndx correspondiente a la carpeta trial de la evaluación del NIST de 2005.....</i>	<i>65</i>
<i>Figura 35. Formato de los vectores cepstrales generados por SPro. Figura obtenida de [SPro Tools]</i>	<i>66</i>
<i>Figura 36. Fichero de salida tras calcular los coeficientes cepstrales a un conjunto de locuciones a través de la función sfbcep de SPro.</i>	<i>67</i>
<i>Figura 37. Ejemplo del formato en el que se presentar los resultados al NIST. Dicho fichero ha sido obtenido en las simulaciones correspondientes a FA.</i>	<i>68</i>
<i>Figura 38. Resultados con respecto a la EER obtenidos en GMM-UBM, con $M=64$ y para los dos modos posibles de SAD.....</i>	<i>72</i>
<i>Figura 39. Resultados con respecto al minDCF($\times 100$) obtenidos en GMM-UBM, con $M=64$ y para los dos modos posibles de SAD.....</i>	<i>73</i>
<i>Figura 40. Valores obtenidos para el minDCF($\times 100$) mediante GMM-UBM sin normalización de puntuaciones y en modo MeanSAD.</i>	<i>75</i>
<i>Figura 41. Representación gráfica de los valores obtenidos para el minDCF($\times 100$) mediante GMM-UBM con $D=50$ y en modo MeanSAD.....</i>	<i>76</i>
<i>Figura 42. Curvas DEC obtenidas mediante GMM-UBM en modo MeanSAD, con $D=50$ y con T-norm.</i>	<i>79</i>
<i>Figura 43. Zoom aplicado a las curvas DEC obtenidas mediante GMM-UBM en modo MeanSAD, con $D=50$ y con T-norm.....</i>	<i>80</i>
<i>Figura 44. Resultados para la EER mediante GMM-UBM en modo MeanSAD, con $M=512$, $D=50$ y con T-norm.</i>	<i>82</i>
<i>Figura 45. Resultados para el minDCF mediante GMM-UBM en modo MeanSAD, con $M=512$, $D=50$ y con T-norm.</i>	<i>82</i>
<i>Figura 46. Rendimiento de GMM-UBM en modo MeanSAD, con $\alpha_{SAD} = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>84</i>
<i>Figura 47. Rendimiento de GMM-UBM para diferentes condiciones de training y test en modo MeanSAD, con $\alpha_{SAD} = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>85</i>

<i>Figura 48. Rendimiento de GMM-UBM y SVM en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>87</i>
<i>Figura 49. Valores obtenidos para la EER mediante NAP en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>88</i>
<i>Figura 50. Valores obtenidos para el minDCF mediante NAP en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>89</i>
<i>Figura 51. Rendimiento de NAP ($K=80$), SVM y GMM-UBM en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>90</i>
<i>Figura 52. Valores obtenidos para la EER mediante FA en modo MeanSAD, $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>92</i>
<i>Figura 53. Valores obtenidos para el minDCF mediante FA en modo MeanSAD, $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>92</i>
<i>Figura 54. Comparación de los rendimientos de GMM-UBM, SVM, NAP ($K=80$), FA-GMM-UBM ($K=90$) y FA-SVM ($K=60$) en modo MeanSAD, $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>93</i>
<i>Figura 55. Fusión entre GMM-UBM y SVM en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>95</i>
<i>Figura 56. Fusión entre NAP y FA en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>96</i>
<i>Figura 57. Comparación de todas las técnicas atendiendo por igual al minDCF y EER en modo MeanSAD, con $\alpha SAD = -0.3$, $M=512$, $D=50$ y con T-norm.</i>	<i>97</i>