BITP3353 Final Exam Notes – Fully Structured

☑ 1. Technology & Applications Related to Software Biometrics

- A. What is Software Biometrics?

Biometrics refers to identifying individuals using physical (e.g., fingerprints, face) or behavioral traits (e.g., gait, voice).

In multimedia databases, biometrics is used for: Authentication and access control, Security monitoring, Indexing and tagging multimedia content.

◇ B. Types of Biometrics

Type	Description	Examples
Physiological	Physical traits	Fingerprint, face, iris
Behavioral	Behavior patterns	Voice, gait, signature
Soft Biometrics	Indirect/non-unique traits (for filtering)	Height, skin tone, clothing

◇ C. Applications in Multimedia Systems

Application Area	Usage
Surveillance	Face/gait detection for public security
Smart Devices	Voice or face authentication
Multimedia Indexing	Tagging people in videos/images

◇ D. Techniques in Software Biometrics

Image Segmentation – separating object from background

Feature Extraction – identifying unique characteristics (e.g., face shape, histogram)

Machine Learning Models – SVM, CNN, HMM for classification/recognition

🧠 Quick Tip: Must-Know Points

Soft biometrics are for filtering, not unique ID

Recognition requires segmentation + feature extraction

Be prepared to explain real-world scenarios (e.g., CCTV, smart login)

☑ 2. SD vs HD (Standard vs High Definition)

- A. What is SD and HD?

Type	Resolution	Common Use
SD	720×480 (NTSC) or 720×576 (PAL)	Older broadcasts, DVDs

HD	1280×720 (720p) or 1920×1080 (1080p)	Streaming, CCTV, YouTube

◇ B. Key Differences

Aspect	SD	HD
Resolution	Lower	Higher
File Size	Smaller	Larger
Bandwidth	Lower streaming needs	Requires more bandwidth
Clarity	Less sharp	Sharper, clearer image

◇ C. Impact on Multimedia Databases

File size: HD files are larger, cost more to store

Query performance: HD videos slow down search/retrieval

Storage: Need for compression tools like H.264, H.265

Bandwidth: Affects streaming quality and speed

🧠 Quick Tip: Must-Know Points

HD = better quality, but heavier storage + slower performance

SD is efficient for archiving, HD is better for analysis or detail

Know compression formats: H.264 commonly used to reduce size

☑ 3. 4Vs of Big Data Applied to Speech Recognition

- A. What are the 4Vs?

V	Meaning	Speech Application Example
Volume	Massive data size	Large voice datasets (training AI)
Velocity	Speed of data input/output	Real-time recognition (e.g., Siri)
Variety	Different forms of input	Various accents, languages, tones
Veracity	Data trustworthiness/accuracy	Misinterpretation due to background noise

◇ B. Use Cases in Speech Recognition

Voice command (e.g., Alexa, Google Assistant)

Automated transcription of meetings

Voice authentication (biometrics)

Multilingual recognition

◇ C. Technologies Involved

ASR (Automatic Speech Recognition)

NLP (Natural Language Processing)

Deep Learning (e.g., CNN, RNN)

Hidden Markov Models (HMM) for sequence prediction

🧠 Quick Tip: Must-Know Points

Volume = large dataset, Veracity = noisy data accuracy

Apply 4Vs to a real system (e.g., Google Assistant, car navigation)

Expect a question that asks how to handle accents/noise

☑ 4. Data Types & Their Impact on System Performance

- A. Common Multimedia Data Types

Type	Use	Stored As
Text	Metadata, captions	VARCHAR, CLOB
Image	Photos, frames	BLOB (e.g., JPEG)
Audio	Sound files	BLOB (e.g., MP3)
Video	Streaming media	BLOB (e.g., MP4)

◇ B. How Data Types Affect Performance

Retrieval time: Text is faster than binary

Storage: BLOBs require more disk space

Indexing: Difficult for binary (audio/video)

Backup & Restore: Takes longer for large multimedia

◇ C. Optimization Techniques

Compression: Use JPEG, MP4, MP3 to reduce file size

Indexing: Content-based indexing for images/videos

Partitioning: Store large media separately from main DB

CDN: Use Content Delivery Networks to reduce load

Caching: Temporarily store frequent media content

🧠 Quick Tip: Must-Know Points

BLOB = binary (image/video/audio), CLOB = large text

Multimedia increases DB load – optimize with compression and indexing

You might be asked to compare tradeoffs (e.g., raw vs. compressed storage)