

**Kurzfassung der Dissertation:**  
**"Human Motion Analysis Based on Organic Computing Principles"**  
(vorgelegt von Dipl.-Ing. Thomas Walther)

Die hochentwickelte Fähigkeit des Menschen, durch bloße Beobachtung Handlungsmuster und Intentionen anderer Individuen zu erfassen, gilt als eine der wichtigsten (non-verbalen) Grundlagen für ein reibungsloses soziales Miteinander in einer komplexen Umwelt. Um diese Fähigkeit zu etablieren, bildet (neben z. B. der Mimik-Analyse) die verlässliche Erkennung und Interpretation von Körperhaltungen und komplexen Bewegungsabläufen (im Folgenden vereinfachend zusammengefasst als *Posenerkennung*) eine unerlässliche Grundlage: künstliche Systeme, welche die menschliche Expertise in diesem Bereich nachahmen, waren und sind Gegenstand intensiver Forschung der computergestützten Bildverarbeitung. Das Anwendungsspektrum marktreifer Posenerkennungssysteme würde sich auf verschiedenste Gebiete erstrecken, einschließlich Medizintechnik, Automatisierungs- und Handhabungstechnik in der industriellen Fertigung, Sicherheits- und Überwachungssystemik und Film- und Sportindustrie.

Jedoch sind, allen Forschungsanstrengungen zum Trotz, die derzeit zur Verfügung stehenden, artifiziellen Posenerkennungs-Lösungen noch immer nicht kompetitiv, verglichen mit ihren biologischen Vorbildern. Eine Erklärung für dieses Defizit findet sich in den gegensätzlichen Paradigmen, die der Informationsaufnahme und -verarbeitung in den vorgenannten Domänen zugrunde liegen: das visuelle System des Menschen lernt während seiner gesamten Existenz selbstständig aus seinen Beobachtungen; die für die jeweilige Aufgabe (hier: Posenerkennung) relevanten Informationen werden selektiv erfasst und in einem komplexen Lernprozess konzeptualisiert. Gelernte Konzepte lassen sich wiederum auf neue Situationen projizieren: diese Fähigkeit zur *Generalisierung* erlaubt z. B. die zuverlässige Analyse bislang unbekannter Szenarien, sowie die Erweiterung vorhandener Konzepte oder deren Neuanlage, sollte kein gespeichertes Konzept auf die neue Situation adaptiert werden können.

Im Gegensatz zum natürlichen Sehapparat lassen künstliche Posenerkennungssysteme derzeit komplexe Lernvorgänge, Ausbildung von Konzepten oder Generalisierungs-Ansätze vermissen: benötigtes Wissen wird weitestgehend vom Entwickler eingebracht, z. B. in Form von geeigneten Körpermodellen und/oder Trainingsdaten. Diese Vorgehensweise führt im allgemeinen zu relativ restriktierten, auf das erwartete Aufgabengebiet spezialisierten Informationsstrukturen, welche sich, aufgrund mangelnder Lern- und Generalisierungskapazitäten, weder dynamisch erweitern, noch sich flexibel an neue Anforderungen anpassen. Als Konsequenz folgt in unbekanntem, komplexen Szenarien eine signifikante Verschlechterung der Systemperformanz, welche im Sinne einer breiten Marktdurchdringung inakzeptabel ist.

Die vorliegende Dissertation stellt, basierend auf den obigen Erkenntnissen, die naheliegende Frage, inwieweit Prinzipien des biologischen Sehsystems für die Verbesserung computergestützter Posenerkennung herangezogen werden können: im Sinne des 'Organic Computing'-Paradigmas wird ein neuartiges Posenerkennungssystem entwickelt, welches, unter Nutzung eines generischen Informationskerns, benötigtes Modellwissen selbstständig aus einfachen Videosequenzen gewinnt. Jedes analysierte Szenario liefert dabei ein szenenspezifisches, auf den Oberkörper beschränktes, 2D-'pictorial structure'-Modell des beobachteten Individuums. Mehrere dieser Modelle werden anschließend in einem abstrakten *meta model* fusioniert, dessen Einzelteile sich während des Fusionsprozesses zu konzeptionellen Repräsentationen der beobachteten Körperteile entwickeln: jeder dieser *Körperteil-Prototypen* kodiert persistente Umriß-, Farb- und Textur-Informationen des von ihm beschriebenen Körperteils; Zusammenhalt zwischen den modellierten Extremitäten wird durch kinematische Zwangsbedingungen gewährleistet, welche ebenfalls vollständig autonom aus den zur Verfügung stehenden Observationen ermittelt werden.

Die sehr guten Generalisierungsfähigkeiten des erstellten Modells werden experimentell in zuvor ungesesehenen Szenarien nachgewiesen: Posen in Standbildern signifikanter Komplexität lassen sich mit hoher Zuverlässigkeit erkennen und könnten, in zukünftigen Evolutionsstufen des Systems, herangezogen werden, um bereits erlernte Meta-Konzepte des menschlichen Körpers weiter zu verbessern. Eine praktische Anwendung im Arbeitsgebiet Robotik rundet die experimentelle Sektion ab: hierbei erhält das erstellte Posenerkennungssystem die Kontrolle über die Kinematik eines humanoiden Roboters (Typ: NAO, Hersteller Aldebaran Robotics), welcher dadurch in die Lage versetzt wird, visuell erfasste Posen eines menschlichen Akteurs nachzuahmen.

Zusammenfassend läßt sich die vorliegende Dissertation als Machbarkeitsstudie verstehen, deren positive Ergebnisse eine weitere Integration von 'Organic Computing'-Prinzipien in konventionelle Posenerkennungs-Lösungen indizieren; der zu erwartende Nutzen und mögliche Synergien mit anderen Disziplinen der computergestützten Bildverarbeitung rechtfertigen die Intensivierung der Forschungstätigkeit in diesem zukunftssträchtigen Bereich.