\* ITALIAN INSTITUTE OF TECHNOLOGY , GENOA (IT).

The core cognitive ability to perceive and synthesize 'shapes' underlies all our basic interactions with the world, be it shaping one's fingers to grasp a ball or shaping one's body while imitating a dance. In this article, we describe our attempts to understand this multifaceted problem by creating a primitive shape perception/synthesis system for the baby humanoid iCub. We specifically deal with the scenario of iCub gradually learning to draw or scribble shapes of gradually increasing complexity, after observing a demonstration by a teacher, by using a series of self evaluations of its performance. Learning to imitate a demonstrated human movement (specifically, visually observed end-effector trajectories of a teacher) can be considered as a special case of the proposed computational machinery. This architecture is based on a loop of transformations that express the embodiment of the mechanism but, at the same time, are characterized by scale invariance and motor equivalence. The following transformations are integrated in the loop: (a) Characterizing in a compact, abstract way the 'shape' of a demonstrated trajectory using a finite set of critical points, derived using catastrophe theory: Abstract Visual Program (AVP); (b) Transforming the AVP into a Concrete Motor Goal (CMG) in
iCub's egocentric space; (c) Learning to synthesize a continuous virtual trajectory similar to the demonstrated shape using the discrete set of critical points defined in CMG; (d) Using the virtual trajectory as an attractor for iCub's internal body model, implemented by the Passive Motion Paradigm which includes a forward and an inverse motor model; (e) Forming an Abstract Motor Program (AMP) by deriving the 'shape' of the self generated movement (forward model output) using the same technique employed for creating the AVP; (f) Comparing the AVP and AMP in order to generate an internal performance score and hence closing the learning loop. The resulting computational framework further combines three crucial streams of learning: (1) motor babbling (self exploration), (2) imitative action learning (social interaction) and (3) mental simulation, to give rise to sensorimotor knowledge that is endowed with seamless compositionality, generalization capability and body-effectors/task independence. The robustness of the computational architecture is demonstrated by means of several experimental trials of gradually increasing complexity using a state of the art humanoid platform.