Proceedings of the Twenty-Third International Florida Artificial Intelligence Research Society Conference (FLAIRS 2010)

Primitive Capabilities for Visual Perception

Sudharsan R Iyengar
Winona State University, Winona, MN, USA
Computer Science Department
siyengar@winona.edu

Abstract

Visual perception is a phenomenon that has puzzled computational scientists for a long time now. Core mnemonics are needed to emulate human cognitive behavior. Mnemonics for aural perception have been developed and designed by the authors. These are used for formulating primitive natural language constructs and for automated musical compositions. We now propose and define fundamental mnemonics that form the basis for Visual Perception. These are a) Present, b) Aligned, c) Rotation, d) Link, and e) Combination, and f) Composition. We present our theory using points and linear arrangements.
Keywords: Mnemonics, Visual Perception, Artificial Intelligence.

Introduction and Fundamentals

Visual perception is a phenomenon that is integral to intelligence [2, 4]. Human visual perception is a phenomenon that has puzzled computational scientists for a long time now. One could use brute force graphics and pattern matching and pattern recognition algorithms and still fall short of the capabilities that a child demonstrates. Our goal is to develop fundamental capabilities and a framework for cognition and learning that can be applied and extended so the system capability and performance can be incrementally improved. We now present a framework of primitives for 2-dimensional, uni-color, symmetric and basic linear structures.

<u>Definition1:</u> A visual element, e, is a unit of information recognized by the system.
<u>Definition 2:</u> A visual element e can be recognized as an atomic element or as an arrangement of known visual elements.
<u>Definition 3:</u> A capability, c, is the application of an operation to an element e, or any sub-element of e.
<u>Definition 4:</u> An element e, is recognized, when e matches with some element in the knowledge base or when the application of capability yields elements that are recognized.
<u>Definition 5:</u> A history is a pair, (V, n) where V is an application of a capability c to an input element e that leads to its recognition and n is the number of times this has occurred.
<u>Definition 6</u>: An element is abstracted and becomes part of the system after a threshold number times it is recognized. An abstracted item atomically recognized, i.e. – the system does not need application of any of its capability to recognize it.
<u>Definition 7:</u> A learning capable system, S= {C, A, H}, is a triple [1], where C is a set of capabilities, A is a set of abstracted elements or knowledge, and H is the history set.

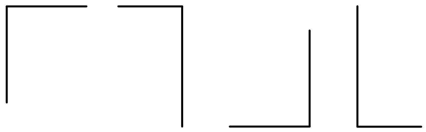
Based on the definitions above, we present the following Visual Capabilities (VC).

VC1: Present (or Absent): A given element, ☐, is recognized as either Present (P) when it is recognized or as Absent (A) – if not recognized. P and A are opposites of each other.
VC2: Aligned: A given element, ☐, is recognized as a linear arrangement of some known element, e.
VC3: Rotation: A given element is rotated to recognize the element.
VC4: Link: A given element is recognized as a juxtaposition of two known elements.
VC5: Combination: Given an input e = e1e2e3…em, where e1, e2, .. em, are recognizable elements themselves, an application of VC5 recognizes e using the following rule:

Given a 2-D plane, we provide four types of arrangements recognized by VC2 – as a linear arrangement of points. Horizontal (H), Vertical (V), Sliding (S), Climbing (C)

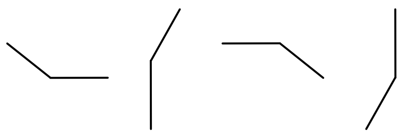
Given two elements the <u>precedence</u> for choosing elements as the first element is as: H, S, C, and V. The linked elements are denoted by ordered sequencing the elements with the spatial relationship between them. We use fixed sized line segments to illustrate our theory.

Copyright © 2010, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.



H BD V; H ED V; H EU V; H BU V

Fig. 1: Linked H and V.



H BU S; C BD V; H ED S; C EU V

Fig. 2: Linked of S or C with H.

Identities and Equivalences:

F(V) > C; *F*(H) S; *F*(S) V; *F*(C) H.
F(P/A) P/A, *N*(P/A) P/A.
N(V) H; *N*(H) V; *N*(S) C; *N*(C) S.
N(ED) EU; *N*(EU) BU;
N(BU) BD; *N*(BD) ED;

F(BU) BU | BD if resulting link has a V.
F(BD) BD | ED if resulting link has a V.
F(ED) ED | EU if resulting link has a V.

F(EU) EU | BU if resulting link has a V.
N(e1 R e2) e1 *N*(R) e2 : {e1, e2 | {H, V}}
N(e1 R e2) e1 *N*(R) e2 : {e1, e2 | {S, C}}
Otherwise *N*(e1 R e2) N(e2) *N*(R) N(e1)
F(e1 R e2) *F*(e1) *F*(R) *F*(e2)

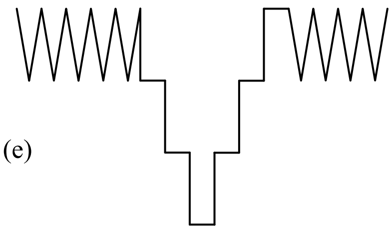
Visual Systems

If a system starts with empty knowledge then no amount of application of any of its capabilities will yield recognition and consequently the system cannot learn. Systems are defined with increasing set of capabilities as below.

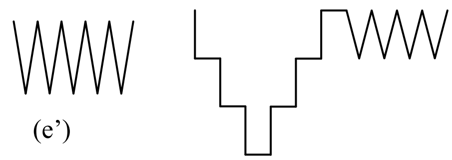
VS1 = {{VC1, VC2}}
VS2 = {{VC1, VC2, VC3}}
VS3 = {{VC1, VC2, VC3, VC4}}
VS4 = {{VC1, VC2, VC3, VC4}}

Once the system is trained, the systems recognizes more complex elements – using content addressable matching abilities – together with an ability to record experiences in its history and ability to abstract/imbibe knowledge from this history of experience.

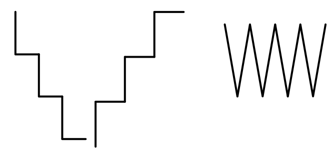
Example 1: Given the input figure below and VS4 = {{., -, |, \, /, L, <, >, V, ^},{}}



(e)



(e')



Using Rule 1 we first recognize this as a combination of
1) an alignment of V (e') and
2) the rest of the input – e".

e" can further be recognized using Rule 1 - alignment of L(e"') - Rule 2 - alignment of V – (e"") and rest of input e""'.
e""' is recognized as an alignment of *N*(L). Thus the entire input is recognized using VC2, VC5 – Rule 1, and Rule 2, VC4 - *N*.

Conclusion:

We have proposed and defined fundamental mnemonic capabilities applicable for visual perception. A methodical application of the capabilities can be shown to enable visual perception. These are similar to mnemonics for aural perception introduced in [1].

References

[1] A Theory on the Abstraction and Cognition Based on Musical Note Patterns, Sudharsan Iyengar, ICAI-04, Las Vegas, June 2004.
[2] A simple visual perception model by adaptive junction, Ajioka, Y. Inoue, K., IJCNN, 1992.
[3] Artificial Intelligence – creation of concepts on Visual Perception's base, Kuznik, V. M., Electronic Journal "Investigated in Russia", 2002.