

Abordagem computacional da percepção da luz.

O termo “percepção da luz” pode ser considerado como uma designação mais ampla que compreende a percepção visual, entre outras possibilidades. Aquilo que torna a percepção da luz como algo que possa ser dito “visual”, consiste em um conjunto de propriedades geométricas que especifica certa classe de invariantes em relação a certos tipos de transformações. Essa proposição traduz, segundo a escola de Erlangen, aquilo que se denomina uma geometria e, desta feita, poderíamos de forma mais particular, denomina-la de geometria visual, ou geometria da visão. A determinação de quais são as transformações adequadas para traduzir a geometria da visão constitui um problema peculiar, cuja investigação é objeto de interesse da ciência da computação e da matemática aplicada e corresponde à área de pesquisa conhecida como visão computacional. Essa disciplina, como típica atividade científica, envolve tanto interesses de ciência fundamental, ao estudar os princípios e propriedades da percepção visual, quanto aqueles de ciência aplicada, ao desenvolver métodos e técnicas para obter visão artificial ou estudar as maneiras como a visão se apresenta nos seres vivos.

A literatura sobre percepção visual nos traz inúmeras obras que apontam vantagens em se adotar aquilo que David Marr e Tomaso Poggio qualificaram como Teoria Computacional da Visão, a qual propõe a investigação da visão segundo três vertentes que, embora possam (e devam) ser conduzidas paralelamente, apresentam certa hierarquia conceitual: (i) a investigação dos princípios computacionais e a consequente elaboração de teorias computacionais baseadas nos mesmos, (ii) a determinação de formas de representação adequada para a informação e a construção de algoritmos eficientes para processá-la de forma abstrata e, (iii) a realização concreta das arquiteturas computacionais artificiais ou a caracterização, tanto anatômica quanto fisiológica, da realização biológica da visão. A essência dessa proposta pode ser sumarizada na ideia de que a percepção constitui uma forma de processamento de informação, que transforma informação obtida através dos órgãos sensoriais em informações apresentadas em formato adequado para a produção do comportamento. Essa concepção original para uma abordagem computacional da visão tem estimulado grande debate através das quatro últimas décadas e alimentado inúmeras controvérsias. Estas essencialmente podem ser agrupadas como argumentos contra uma visão mecanicista de aspectos da mente, no presente caso, da percepção, os quais sugerem que os processos biológicos envolvidos na percepção poderiam ser entendidos como manipulações de símbolos. De fato, muito se fez sob esse paradigma, dito “computacionalista”, sendo que os tais “símbolos” que são manipulados nesses processamentos foram evoluindo para entidades cada vez mais próximas do que poderia ser passível de constituir correlatos biológicos dessas entidades abstratas.

O advento do conexionismo trouxe novas possibilidades de se conceber o processamento da informação sensorial, sem utilizar representações simbólicas no sentido lexical típico do computacionalismo, lançando mão de representações baseadas em invariantes dinâmicos ou estatísticos. Os métodos tradicionais de reconhecimento estatístico de padrões baseiam-se em atributos (“*features*”) que constituem de fato construtos locais cujas propriedades apresentam baixa variabilidade estatística em relação a certas classes abstratas, e o processo de inferir essas classes convencionou-se denominar de técnicas de classificação. A disciplina que emergiu da fusão das

técnicas estatísticas de inferência e do emprego de abordagens conexionistas passou a ser denominado de aprendizagem de máquina (“*Machine learning*”).

A aprendizagem de máquina evoluiu ao longo das décadas, explorando métodos ditos não supervisionados, os quais apresentam maior aplicabilidade à construção de modelos que possivelmente possam ser considerados como explicações de processos que se desenvolvem na percepção visual encontrada no mundo animal. A última novidade nesse tipo de abordagem constitui a chamada “aprendizagem profunda” (traduzindo-se diretamente de “*deep learning*”). Esta abordagem emprega métodos de agrupamento hierárquico de atributos filtráveis, os quais têm produzidos resultados de boa qualidade com bom desempenho. Todavia, os métodos agrupados sob o título de aprendizagem de máquina são essencialmente *ad hoc* e nota-se a falta de uma teoria mais fundamental que explique e determine quem são os atributos (features) e por que, como eles devem ser agrupados, quais suas relações hierárquicas e assim por diante. A abordagem da aprendizagem de máquina é essencialmente algorítmica – falta-lhe uma teoria computacional.

O que se pretende discutir neste trabalho é como poderia empreender-se uma teoria computacional da visão sem retornar à ótica mecanicista do computacionalismo e sem fundamentar-se em uma abordagem conexionista puramente *ad hoc*. Curiosamente, as sementes para tal proposta encontram-se nas próprias abordagens originais de Marr e Poggio e passam pela trajetória construída ao longo dos anos por Tomaso Poggio e seus colaboradores. A ideia consiste em partir do estudo do problema da visão, assim chamando-se o conjunto de desafios encontrados pelo processo evolutivo das espécies animais, que conduziu às diversas “soluções biológicas para o problema”. Esse problema da visão caracteriza-se de maneira pragmática: o contexto e o uso são os ingredientes requeridos para se caracterizar as propriedades invariantes em relação ao conjunto de transformações que compõem o processamento de informação da percepção visual.